THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 1996, *49A* (3), 525–545

A General Factor Involved in Dual-task Performance Decrement

Patrick A. Bourke
School of Psychology, University of Luton, Luton, Bedfordshire, U.K.
John Duncan and Ian Nimmo-Smith
MRC Applied Psychology Unit, Cambridge, U.K.

Beyond specific conflicts between tasks that are obviously similar (e.g. two verbal tasks) and limits specific to speeded responses, is there a general limitation on what tasks can be done simultaneously? In two experiments, we examined dual-task combinations designed to avoid known sources of specific interference. Under these circumstances, a general factor model predicts consistency in the pattern of results. Tasks should be ordered in demands on the general factor as measured by interference with concurrent tasks; this order should be the same for any concurrent task used to measure it. This prediction was confirmed in both experiments, each involving 12 dual task combinations of four tasks. In Experiment 1, the tasks were tone discrimination, random letter generation, a manual–tactile manipulation task, and recognition memory for photographs. In Experiment 2, the first of these was replaced by an easier tone-monitoring task, and the last by a visual prototype learning task.

There are many instances of interference between concurrent but quite different tasks. Several theorists have suggested that this may be due to a common underlying factor. Broadbent (1958, 1982) has argued that there is a general limitation on how much information the entire cognitive system can transmit at one time. Moray (1967) and Kahneman (1973) have proposed that there is a general resource or processor that must be drawn upon for the successful performance of almost all tasks. Baddeley and Hitch (1974) proposed a central executive that monitored and coordinated performance of a wide range of tasks. In different contexts similar proposals have been made by many others (e.g. Ackerman, Schneider, & Wickens, 1984; Duncan, Williams, Nimmo-Smith, & Brown, 1993; Norman & Shallice, 1980; Posner, 1978; Yee, Hunt, & Pellegrino, 1991). In all variants of this view, the common factor is considered to be limited in its ability to support multiple tasks simultaneously. Dual-task interference is thought to occur when concurrent demands on the common factor are too great to be met.

Requests for reprints should be sent to Patrick Bourke, School of Psychology, University of Luton, Luton, Bedfordshire LU1 3JU, U.K. Email: patrick.bourke@luton.ac.uk

We wish to thank Michael Posner, Steve Keele, Jon Driver and anonymous reviewers for their comments on earlier versions of this paper.

An alternative position argues that different pairs of tasks interfere for different specific reasons. For example, Treisman and Davies (1973) have shown that a shared input modality increases dual-task interference. Pashler (1990) has shown substantial interference when two tasks require similar speeded motor responses and the order of responding is unpredictable. Hirst and Kalmar (1987) have shown that similarity of semantic domain can be important. Other known sources of specific interference include concurrent phonological coding (e.g. Salame & Baddeley, 1982), common spatial encoding (e.g. Baddeley & Lieberman, 1980), and common visual encoding (e.g. Neisser & Becklen, 1975).

Pashler (1990) has argued that there is specific interference when tasks require the same stage of processing. If two tasks simultaneously require response selection, one will have to queue, and the response to the other task will be delayed. This is thought to apply to many task combinations.

However, these particular sources of specific interference are insufficient to account for all findings of dual-task interference. For example, shadowing interferes with covert visual orienting (Posner, Sandson, Dhawan & Shulman, 1989), silent reading interferes with tone detection (Eriksen & Johnson, 1964), card sorting interferes with random letter generation (Murdock, 1965), mental arithmetic interferes with perception (Reisberg, 1983), and responding to light stimuli can lead to a failure to detect auditory events (Colavita & Weisberg, 1979). How are such results to be explained?

Allport (1980) argues that when two tasks that apparently have nothing in common interfere with each other, it may be for a specific reason that we have not yet identified. In a similar vein, Navon (1984) argues that performing any task will produce side-effects. So, even if two tasks have nothing specific in common, the side-effects of one may interfere with the other. These arguments are useful in that they show that no compelling evidence exists for a common general limitation on dual-task performance. It remains an open question, therefore, whether in addition to the many specific interference effects that have been shown, there is a further, more general source of dual-task interference.

Despite this lack of resolution, the debate has all but died in recent years. In some cases Navon's (1984) criticisms go unaddressed, and results continue to be interpreted in loose general-resource terms (e.g. Just & Carpenter, 1992). At the opposite extreme, the idea of modularity in cognitive operations has become so popular that any possibility of a "general resource" may seem entirely implausible.

The question can be investigated by dual-task studies that avoid the known sources of specific interference. If all or most remaining interference is due to demands on a single common factor or resource, then across a range of related dual-task combinations a particular, consistent pattern of decrement can be predicted. We consider this pattern next.

The Predicted Pattern of Secondary-task Performance

Following Navon and Gopher (1979), let us characterize the general factor or resource hypothetically as: (1) limited, (2) of a fixed amount, (3) split entirely between the two tasks being studied, and (4) producing improved performance as its involvement in a given task increases. We can expect different tasks to place different demands on the general factor.

The ordering of these demands can be assessed by the extent to which the different tasks interfere with a fixed concurrent task, when specific sources of interference are avoided. Different assessments of demand can be made by using different concurrent tasks. If the only reason for interference is concurrent demand for the general factor, then we expect the independently assessed orderings of demand to be the same.

This prediction is represented in Table 1. The table assumes a design in which subjects are told always to treat one task of a pair as "primary"—i.e. to do it as well as possible—and the other as "secondary"—i.e. to do it as well as possible given the constraint of maximizing primary performance. Four secondary tasks (indicated by lower-case letters) are combined in turn with four primary tasks (indicated by upper-case letters). Assume that the general factor demand of the four primary tasks decreases in the order A to D. Table 1 shows the prediction that the most demanding Primary Task A will interfere more than Primary Task B with all four secondary tasks. Similarly, Primary Task B will always interfere more than C, and C will always interfere more than D.

The experimental design that is used in these studies follows the logic outlined above but is extended by using the same four tasks as primary and as secondary (Table 2). This is indicated by the use of the same letters in upper- and lower-case script. Each of the four tasks is treated as primary three times as it is combined with each of the other three tasks treated as secondary. Each of the four tasks is treated as secondary three times as it is combined with each of the other three tasks treated as primary. In addition, single-task performance is tested for each of the four tasks. These single-task cases are represented by the blank cells in Table 2.

Consider first the simple case where subjects maintain a constant level of performance on the primary task (primary task protection). To enable this, each primary task makes some fixed demand on the general factor. The greater the demand that a primary task places on the general factor, the greater will be the decrement observed on the accompanying secondary task. For example, Primary Task A is seen to produce more interference on Secondary Task b than does Primary Task C. This reflects the fact that Task A places more demand on the general factor than does Task C. Given this, the effect must be seen again in the amount of interference observed in Secondary Task d when done with Primary Tasks A and C. Again, Primary Task A must produce more interference

TABLE 1
Predicted Interference of Four Primary Tasks with Four Secondary Tasks According to the General Factor Hypothesis

Primary Tasks	*Secondary Tasks* *w*	*x*	*y*	*z*
A	4	4	4	4
B	3	3	3	3
C	2	2	2	2
D	1	1	1	1

Note: 1 to 4 indicates increasing decrement in secondary task performance.

TABLE 2
The Experimental Design and Predictions with Four Tasks Used as Primary and Secondary Tasks

Primary Tasks	*Secondary Tasks* *a*	*b*	*c*	*d*
A	—	3	3	3
B	3	—	2	2
C	2	2	—	1
D	1	1	1	—

Note: 1 to 3 indicates increasing decrement in secondary task performance.

than Primary Task C. With the current design, such comparisons can be made directly with six pairs. Furthermore, we can make indirect inferences about the order we should expect. For example, if we find that A interferes more than B and B interferes more than C, we must expect A to interfere more than C. All these predictions are summarized in Table 2. It shows that the four independently obtained orders of primary task demand, as measured by performance on the secondary tasks, should be consistent with one another.

In the above argument, we have assumed that there is primary-task protection. However, primary-task protection is rarely achieved. As noted by Kahneman (1973) and Navon and Gopher (1979), when instructions to maintain primary-task performance are given, they are rarely successfully followed. Sometimes this primary-task decrement is large (Tune, 1964, Murdock, 1965), sometimes it is small (Shallice, McLeod, & Lewis, 1985), but it is almost always there. To deal with this, we need a specific model of how the general factor is assigned in a dual task experiment—the issue termed by Kahneman (1973) "allocation policy".

The Weighting Model

We propose that, in a dual-task experiment, each task is assigned a weight, indicating how strongly it competes for the limited general factor. Allocation of the general factor to a task increases with its weight but decreases with the weight of the other, competing task. Many factors could determine the weight attached to a given task, such as subjects' ideas of perfect performance or subjects' ideas of acceptable performance, in addition to the objective function relating performance to general factor allocation. The model makes no assumptions as to what these factors are, except to assume that, for a given instruction to treat that task as primary or secondary, the weight assigned as a result of that instruction remains constant. A (primary or secondary) task competing strongly with any one concurrent task will also compete strongly with all others.

One factor affecting weight is the instruction to treat a task as primary or secondary. This effect will be considered separately from the composite effect of other factors on weight, which for convenience will be grouped under the term "intrinsic weight".

Now consider again the design shown in Table 2. Weights of the four tasks are determined by the two factors distinguished above—intrinsic weights and priority instruction. If intrinsic weights decrease in the order A to D, then we expect this to be the order of overall weights both for the four tasks treated as primary and for the four tasks treated as secondary. The instruction to treat tasks as primary rather than secondary increases all of the weights but does not change the overall order.

Predictions for secondary-task performance may now be derived as before. The greater the weight of a primary task, the greater will be its interference with all concurrent secondary tasks. As before, the four orders of interference established with Secondary Tasks a to d should be mutually consistent.

A new parallel prediction can now be made for the order of decrement on primary tasks. The order of interference with each primary task should again be determined by the decreasing order of intrinsic weights a to d. As the same tasks (with the same order of intrinsic weights) are used as primary and as secondary tasks, the same order of decrement should be seen within the primary tasks as is seen within the secondary tasks. In

Table 2, *primary* task decrement should decrease regularly from left to right along each row.

Single- to Dual-task Decrements

The most conventional way to investigate dual-task interference is analysis of decrements from single-task performance. Does the model we have outlined make any prediction concerning sizes of decrement on the different tasks employed?

A traditional approach to this problem is provided by the framework of Performance Resource Functions (PRFs; Norman & Bobrow, 1975). PRFs map the relationship between performance achieved and extent of general resources hypothetically allocated. In general, the PRF is a *theoretical* function, whose shape in practice is unknown. Two possible PRFs (a) and (b) are shown in Figure 1.

On the abscissae, the general resource allocation is shown increasing from left to right. On the ordinates, task performance is shown improving from bottom to top. Typically, as fewer resources are allocated, worse performance results. The shape of PRFs may in general be assumed to vary with different tasks. In the examples shown (a) has a shallow slope—i.e. as the level of resource is decreased, only a small gradual decrease is seen in performance (until very low resource allocation). In contrast, (b) shows a steep PRF function—i.e. as the level of resource is decreased, a large decrement is seen in performance.

Given two such PRFs, one can reasonably argue that the same small reduction in resource allocation to both tasks would produce little percentage decrement in Case a but a large percentage decrement in Case b. This reduction may happen, for example, when a second task must be done at the same time. In general, the steeper the PRF, the more sensitive will that task be to the withdrawal of the general resource, and the greater the performance decrement that will result.

To predict in practice which tasks will produce the greatest dual-task decrements, however, one needs to know both the shape of PRFs and the weighting rules determining how the general factor is allocated. Consider concurrent performance of the two tasks

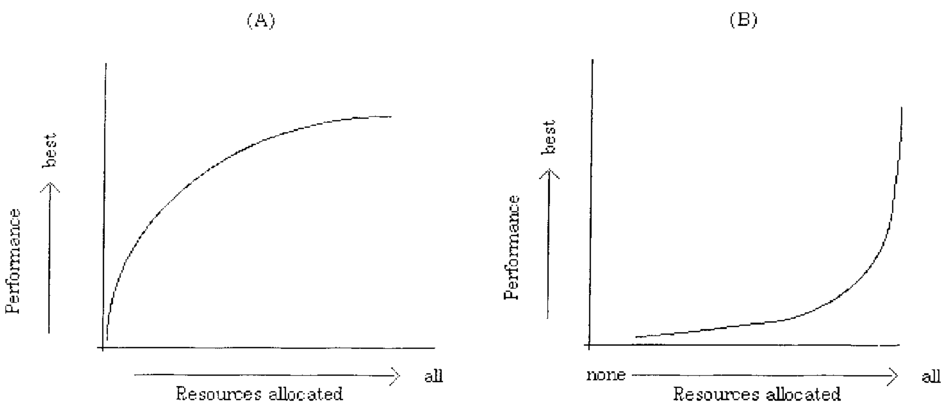

FIG. 1. Two hypothetical Performance Resource Functions (PRFs).

whose PRFs are shown in Figure 1. What can be predicted about the relative sizes of dual-task performance decrement? If the general factor is split evenly between the two tasks, then (a) would show little decrement from single-task performance, but (b) would show a large performance decrement. If, however, (a) received little of the limited general factor and (b) received virtually all of it, then we might find little decrement on (b). This latter could, for example, occur if a task that needs more resources has more resources allocated to it—a common assumption in dual-task research (Navon & Gopher, 1979). In any case, as PRFs are in general unknown, no specific predictions about relative decrements in different tasks can be made. In contrast, the only assumption about PRFs needed to support our approach is that they should always be monotonically increasing. Though single- to dual-task decrements will be reported for completeness, we shall have little further to say about them.

EXPERIMENT 1

A tone discrimination task, a random letter generation task, a manual tactile task, and a visual recognition task were chosen as a diverse group of four tasks that would sample cognition widely. The only constraint in choosing the tasks was an attempt to avoid known specific sources of interference. We avoided tasks that shared a common input modality, common mode of response, shared linguistic coding, shared spatial coding, common semantic domains, or the response selection bottleneck as identified by Pashler (1990).

Method

Subjects

Subjects were 16 paid volunteers. Their modal age was 21, and they were typically recent graduates. There were seven males and nine females.

Design

On each day, 16 conditions were tested: 12 dual-task conditions and 4 single-task conditions. Orthogonal combination of 4 primary with 4 secondary tasks would yield 16 dual-task conditions; when the design called for a task paired with itself, subjects did a single-task condition instead.

On each of three days, subjects had one 3-min trial per condition. Data from the first day were discarded. On each day, all conditions with a given primary task were done together in a block. Blocks were counterbalanced across subjects on each day, and for any given subject were run in opposite order on Days 2 and 3. Within each block, the order of secondary tasks was fixed for a given subject and day, but again was counterbalanced across subjects on each day and reversed for each subject between Days 2 and 3. Sessions lasted about two hours, with a break midway through.

Tasks

Tone Task. This required discrimination between a target tone (approximately 520 Hz, 73 db) and a background tone (approximately 470 Hz, 73 db). The task was to press a foot switch whenever a target tone was detected. Examples of target and background tones were played to the subject immediately before each tone task trial.

Tones were presented over headphones under the control of a BBC microcomputer. Each tone lasted 250 msec. Onset-to-onset intervals between tones varied randomly between 1.56 and 3.70 sec, with a median interval of approximately 2.25 sec. Through an oversight, this meant that sometimes the whole sequence of 58 tones could last a few seconds more or less than 3 min. The tone task was always scored for its actual duration, whereas the task with which it was combined was scored for exactly 3 min. Approximately 50% of tones were targets.

To minimize possible response interference (Pashler, 1990) with other tasks, subjects were told to press the foot switch at their leisure. Responses were automatically recorded and scored by the BBC microcomputer. Responses were always considered to have been made to the immediately preceding tone in the sequence, and only one response (the first) per interval between tones was recorded. The score was simply the total number of correct responses (press to target, no press to background) out of 58. When this task was to be treated as primary, subjects were told their score at the end of the trial.

Random Letter Generation. In this task letter names were to be spoken out loud, in random order, for 3 min. The instructions to subjects emphasized that they should try to be as random as possible, that they should not spell words, use initials, or repeat sequences. Subjects were allowed to generate letters at their own pace. In previous experiments (e.g. Baddeley, 1966), subjects have been asked to generate letters in rhythm with a metronome. In pilot work, however, we found that subjects experienced great difficulty treating random letter generation as secondary with external pacing. Accordingly, subjects were told to generate letters at as fast a rate as they felt was comfortable while consistent with randomness. They were told that they could change the rate at which they generated letters from moment to moment. Occasionally, when the task was first explained, subjects attempted to generate items at a rate that could not be recorded because of poor articulation. In this case they were asked not to go quite so fast.

The experimenter recorded the letters produced by keyboard entry as they were spoken. Letters were also recorded on cassette. If any doubt was felt as to the accuracy of the keyboard entry, data were re-entered from the cassette recording after the session. Randomness was scored by the standard information theoretic measure of information generated (H), based on single element frequencies, with the Miller–Madow correction for bias as described by Baddeley (1966). A measure of how equally the different possible responses are used, H ranges from 0.00 when a sequence is composed entirely of repetitions of a single letter to 4.70 when each letter occurs an equal number of times. Previous work (e.g. Baddeley, 1966) has shown that other measures of randomness (e.g. based on diagram frequency) give similar results.

The initial analysis used simple H values, but we also wished to correct for number generated, as it is well known that letter sequences become less random as the rate at which they are generated increases. This correction was calculated on the basis of data reported by Baddeley (1966). In Baddeley's (1966) experiment, which used subjects similar to ours in age and background, letters were generated at rates from 0.5 to 4 sec per letter in time with a metronome. From the maximum possible value of 4.70, almost achieved at the slowest rate, results showed a decrease in H of 0.063 for each doubling of speed. In our analysis, accordingly, each H value for each subject was adjusted by an amount corresponding to 0.063 times the extent to which the corresponding number generated exceeded, in $\log_2$ or doubling units, the harmonic mean for that subject across days and conditions. Analyses based on these corrected scores are reported here; however, uncorrected scores gave much the same results.

As with the other tasks, when random generation was to be treated as primary, subjects were usually told their score (uncorrected H and number generated) after each trial. They were told that H was the main score, explained simply as a sensitive measure of randomness. However, no immediate feedback was given on the manual/random-generation combination. Here, the experimenter was recording performance on the manual task and could not input the letters. Accordingly, subjects were told their score at the next experimental session.

Manual Task. Two bolts were mounted upright on a wooden base. The task was to screw a nut first down to the bottom of one bolt and back up to the top, then down to the bottom of the other bolt and back up to the top, repeating this sequence continuously and as quickly as possible for 3 min.

The base on which the bolts were mounted was enclosed on two sides and on the top. The top and side sections were to ensure that the subjects did not use visual information in performing the task. The front opening allowed the subject access to the nut and bolts. The sides, top, and base of the apparatus were made of wood. The base was 25 cm square and 2 cm in depth. The walls were 13 cm high towards the front and 22 cm towards the back, producing a sloping top. An aluminium sheet was attached to the base. Mounted on this were the two plastic bolts, each 5 cm in height and 1 cm in diameter. The bolts were mounted on a line 18 cm from the front. They were both 9 cm from the side walls and separated from one another by 5 cm. The nut was made of Perspex, 3 cm squared (approx.), and made to fit the thread of the bolt closely.

Each trial started with the nut on the top of the bolt on the subject's right-hand side. Performance was recorded by the experimenter, who sat opposite the subject in any condition that involved the manual task. Each time the nut was brought to the bottom of either bolt, it was scored as one. If the subject had brought the nut to the top of a bolt when 3 min had elapsed, this was scored as an additional 0.25. If the subject had brought the nut across to the other bolt, this was counted as 0.75.

Visual Recognition. This task had two phases—learning and test. In the 3-min learning phase, subjects watched a video-recorded series of still coloured photographs taken from magazines. The test phase, in which recognition of the previously presented pictures was assessed, followed immediately. In dual-task conditions, only the learning phase was carried out concurrently with other tasks.

Seven 3-min sequences were composed. Each sequence consisted of 120 photographs, which had been videotaped and edited together so that each one lasted for 1.5 sec. No photograph was repeated. All photographs filled the monitor screen, which was approximately 41 cm $\times$ 32 cm in size. Subjects sat about 63 cm from the screen. Each sequence was used once per day in one of the seven conditions involving visual recognition.

For the recognition phase, three test tapes of 30 photographs each were prepared for each of the seven learning tapes. Different recognition tapes were used each day. In each recognition tape, half of the photographs were new, and half were from the relevant learning tape. The subject was to decide which were old and which were new and indicate this decision by saying "new" or "old". During the recognition test, subjects were allowed as much viewing time as they desired. Responses were recorded by the experimenter.

Photographs were randomized as follows. First they were collected and sorted into categories: faces, human figures, animals, objects, scenery, interiors, exteriors, miscellaneous. Approximately equal numbers from each category, amounting to 120 photographs, were placed in each of eight piles. Seven were used to form the seven learning tapes. Photographs in each of these piles were videotaped, in random order, to produce seven 3-min sequences. The photographs in the eighth pile were used as "new" items when the recognition sequences were constructed.

For each of the seven learning tapes, the number of items from each category occurring in each minute of the 3-min presentation was counted. A corresponding proportion of these categories was used in each third of the corresponding test tapes. The old items used in the recognition sequence were presented in the same order as in the learning sequence. The same number of photographs from each category were used to make up the new items. Old and new items were interspersed randomly through the sequence.

The measure of visual recognition was simply the number of correct "old" and "new" responses to a 30-item test tape. When the recognition task was to be treated as primary, subjects were told their score at the end of the trial.

Instructions

Subjects were told that the experiment was looking at how much people could attend to at one time. After the tasks were explained and demonstrated, it was emphasized that there was one final, very important point. This was that in each combination one task would be designated the main task and the other the secondary task. They were told that they should try to do as well as possible on the main task and then, given that constraint, as well as they could on the secondary task. Before each trial, the subject was reminded of which task was to be treated as the primary or main task.

Significance Testing

There are limitations to the usefulness of conventional analysis of variance (ANOVA) and multiple comparison procedures for determining whether means not only differ from each other, but do so in a way that is consistent with an order restriction. An appropriate analysis will be one that is in part analogous to adopting a one-tailed as opposed to a two-tailed test for the difference of two means when there is a prediction of the direction of any effect.

Isotonic regression (*IR*) is an established method for doing this in the case where there is a set of order predictions concerning the means of one performance measure (Robertson, Wright, & Dykstra, 1988). *IR* takes a family of means on one performance measure and a hypothesis concerning the order of those means. *IR* then determines the minimal change to the data that would be required to make it consistent with the hypothesized order and then adjusts these required changes by an estimate of the subject by task variance component of the data. The significance of the required scaled changes to the orginal data (the S12 statistic) is assessed by reference to a table of critical values (Robertson et al., 1988, Table 6). S12 measures the extent and significance of deviations from the predicted order. If S12 is significantly large, then the set of means is not consistent with the hypothesized order.

Furthermore, *IR* provides a complementary statistic, S01, which is used to test whether there are significant differences between the means in so far as they are *compatible* with the hypothesized order. The significance of S01 is assessed by reference to a second table of critical values (Robertson et al., 1988, Table 7).

There is an analogy with linear regression: S01 corresponds to the test of whether the slope of a linear regression is non-zero; S12 corresponds to the test for non-linearity of the regression (e.g. the presence of quadratic or higher-order components).

The nature of the present design means that a further extension has to be made to standard *IR*. The first point is that more than one set of means is involved, concerning more than one performance measure. In fact, each analysis involves 8 sets of 3 means (four tasks, each occurring as both primary and secondary). The second point is that the set of order predictions is not determined in advance. This is because we do not specify a priori the relative extent of demand imposed by any of the tasks. Instead, the hypothesis being considered is that, in so far as there are differences amongst the means of the measures (i.e. differences in general demands of the primary and/or secondary tasks), the directions of such differences should be consistent across the various primary and secondary tasks.

The requirements of this form of data and hypothesis have been met in this study by the following extension of standard *IR*, which we call *Consistent Isotonic Regression* (*CIR*). All possible orders of demand of the four tasks are considered. A "best order" is then chosen as the one that minimizes the aggregate inconsistency between any given order and all the sets of means. (The aggregated inconsistency is the sum of the S12 statistics for the participating sets of means; this is equivalent to the weighted total across the various performance measures of the sums of squares of deviations of the means from the fitted isotonic values, the weights being inversely related to the estimated variances of the corresponding performance measures.) This "best order" is then used in the same way as a hypothesized order in standard *IR*. For each of the (eight) sets of means, the individual S12 statistics are computed relative to the "best order". The significance of each of these S12 statistics is assessed individually, allowing for the possibility that each set of means might contain effects that differed significantly from the "best order". The complementary statistics S01 are assessed in parallel to test whether there is indeed evidence of significant effects consistent with the "best order".

This procedure might be criticized because the "best order" has been chosen precisely to capitalize on whatever consistency there is between the various sets of means. An alternative "cross-validation" approach is available—that is, to drop the set of data from each task in turn; to determine the "best order" from the remaining sets, and then to use standard *IR* to test the consistency of the excluded set with this "best order". There is a conceptual disadvantage of this cross-validated *IR* in comparison with *CIR* in that at the end there might not necessarily be a single order in view. In the present studies there is a high level of consistency between *CIR* and this cross-validated *IR*. The single exception will be noted later.

Results

As estimated by *CIR* analysis (see further on), the estimated "best order" of tasks—in order of decreasing general factor demand—was: tones, random generation, visual recognition, manual. In the data set as a whole, in other words, "tones" produced the most interference with other concurrent tasks, and "manual" produced the least.

The extent to which this single order was consistent with data from each individual task is shown by the mean data in Figure 2. Each panel shows data from one task, carried out singly (crosses), as a primary task (solid dots), and as a secondary task (open dots). Concurrent tasks are shown along the abscissa, with single-task performance at the right. In each panel, the three concurrent tasks have been ordered consistently with the overall "best order" given above. According to our hypothesis, therefore, we should see a consistent *increase* in scores from left to right in each panel, as the demand of the concurrent task *decreases*. We might also expect to see primary performance better than secondary, and best performance in the single task.

Mean performance on the tone task is shown in the top left panel of Figure 2. Results are approximately consistent with the overall "best order" in that, from visual inspection, concurrent random generation produces worse performance than concurrent visual recognition or manual tasks; the comparison between the latter two, however, shows a small violation of the overall "best order". Single-task performance is best. As expected, primary performance is better than secondary.

Mean performance on random generation is shown in the top right panel. Again from visual inspection of the figure, results are consistent with the overall best order in that the concurrent tone task produces worse performance than the visual task, which, in turn, produces worse performance than the manual task. Single-task performance is best. Contrary to expectation, there is little difference between primary- and secondary-task performance.

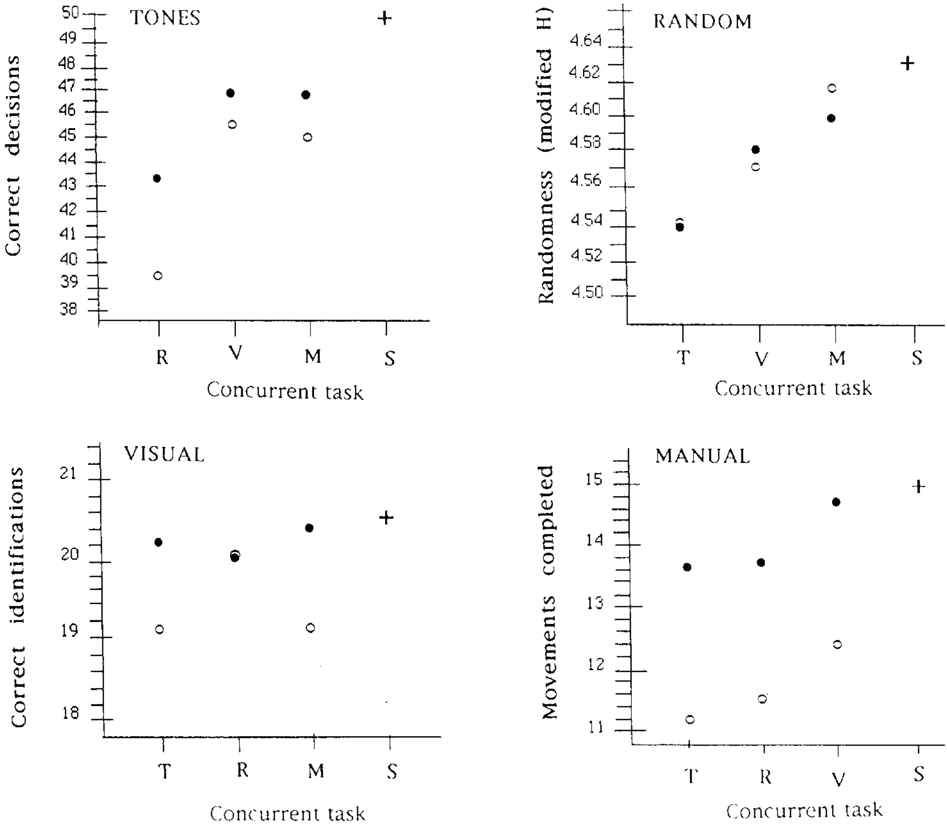

FIG. 2. Experiment 1: Average performance in all tasks and conditions. T = Tones, R = Random, V = Visual, M = Manual, S = Single Task. + = Single task score, ● = Primary task score; ○ = Secondary task score.

Mean performance on the visual recognition task is shown in the bottom left panel. Results, from visual inspection, are inconsistent with the overall best order. Single-task performance is best. Primary performance is generally better than secondary.

Mean performance on the manual task is shown in the bottom right panel. Results are consistent with the overall best order in that the concurrent tone task produces worse performance than the random generation task, which, in turn, produces worse performance than the visual task. Single-task performance is best. Primary performance is better than secondary.

Statistical analysis (the *CIR* analysis) of these results is shown in Table 3. The analysis indicates the extent to which each of 8 sets of 3 means shows variation consistent with a single overall "best order". The 8 sets arise from each of the 4 tasks, occurring as both primary and secondary, being performed with the remaining 3 tasks. For each task, an average error term from primary and secondary conditions has been used. Single-task data do not enter the analysis. As stated above, the obtained "best order"—the order best fitting the whole of the data—was: tones, random generation, visual recognition, manual (decreasing general factor demand). Table 3 shows the extent to which data from each set

TABLE 3
Isotonic Regressions Assessing Goodness of Fit of the Eight Orders of Decrement to a Single Best Order

Tasks		*SS between Tasks*	*Variance Explained*	*S01*	*p*	*Variance Not Explained*	*S12*	*p*
Tones	P	134.63	134.43	6.72	<.05	0.21	0.01	n.s.
	S	351.50	350.22	17.50	<.01	1.28	0.06	n.s.
Random	P	0.03	0.03	2.39	n.s.	0.00	0.00	n.s.
	S	0.06	0.06	5.37	<.05	0.00	0.00	n.s.
Visual	P	0.73	0.47	0.17	n.s.	0.26	0.10	n.s.
	S	10.67	2.67	0.98	n.s.	8.00	2.95	n.s.
Manual	P	14.83	14.83	15.42	<.01	0.00	0.00	n.s.
	S	12.99	12.99	13.51	<.01	0.00	0.00	n.s.

Notes: P = primary task, S = secondary task. Critical values—S01 for 3,60 observations: .10 = 2.62, .05 = 3.93, .01 = 7.20. Critical value—S12 for 3,60 observations: .10 = 3.38. S01 assesses consistency, and S12 inconsistency, with the best order.

of means showed variation consistent (S01) or inconsistent (S12) with this overall "best order".

For 3 of the 8 sets of means, there was no significant variation, either consistent or inconsistent, with the "best order". In particular, the visual recognition task—whose data seemed most discrepant from the overall pattern (Figure 2)—in fact showed no significant effects of concurrent task. For the remaining 5 sets of means, there was significant between-condition variability, which was all consistent with the overall "best order".[1]

For each task, an ANOVA was used to assess the effect of dual-task conditions—the task when done alone, done as a primary task (averaged over the three levels), and done as a secondary task (averaged over the three levels). The effect was as follows: tones, $F(2, 30) = 15.02$, $p < .0001$; random generation, $F(2, 30) = 2.53$, n.s.; visual recognition, $F(2, 30) = 1.13$, n.s.; manual, $F(2, 30) = 21.70$, $p < .0001$. Three planned contrasts were performed on each task: single compared to primary, single compared to secondary, and primary compared to secondary. The tone task was performed better as a single task than as a primary task, $F(1, 30) = 9.00$, $p < .005$, and better as a single task than as a secondary task, $F(1, 30) = 29.95$, $p < .0001$. It was also performed better as a primary task than as a secondary task, $F(1, 30) = 6.11$, $p < .02$. The random generation task was performed better alone than when performed as a primary task, $F(1, 30) = 4.46$, $p < .04$, but there was no reliable difference between when it was performed as a single task

[1] With only a single exception, cross-validated *IR* (see method) gave exactly the same results as the overall *CIR* analysis. The exception concerned analysis of random generation scores. Excluding random generation data changed the estimated best order from tones, random generation, visual recognition, manual to tones, random generation, manual, visual recognition. However, assessing the fit of random generation scores to the latter order—obtained without taking them into account—again showed no significant violations.

and when it was performed as a secondary task, $F(1, 30) = 2.99$, $p < .09$. There was also no difference between when it was performed as a primary task and when it was performed as a secondary task, $F(1, 30) = 0.147$, n.s. The visual recognition task was performed no better when done alone than when performed as a primary task, $F(1, 30) = 0.147$, n.s., or as a secondary task, $F(1, 30) = 2.11$, n.s. There was also no difference between when it was performed as a primary task and when it was performed as a secondary task, $F(1, 30) = 1.15$, n.s. Performance on the manual task when done alone compared to when done as a primary task just failed to be significantly different, $F(1, 30) = 4.00$, $p < .055$. There was a difference between when it was performed alone and when it was performed as a secondary task, $F(1, 30) = 41.42$, $p < .0001$. There was also a difference between when the manual task was performed as a primary task and when it was performed as a secondary task, $F(1, 30) = 19.68$, $p < .0001$.

Discussion

To the extent that tasks showed significant variability due to the concurrent task with which they were paired, all such variability could be explained by assuming a single underlying order of general factor demand. The results are consistent with the view that, once known sources of specific interference are avoided, each task is characterized by an overall "demand" reflecting how strongly it will interfere with *any* dissimilar concurrent task.

There are, however, weaknesses in the experiment. First is the lack of significant variation in 3 of the 8 sets of means examined. The visual recognition task may simply have been insensitive. It is possible that too many pictures were presented too fast, so that irrespective of how much of the general factor was invested, little significant improvement could occur. In Experiment 2, this task was replaced with an alternative that still satisfied the requirement of avoiding specific similarity to other tasks. This new task required the subject to watch a series of exemplar patterns for 3 min and then to draw the prototype pattern from which they thought the patterns were derived (cf. Posner & Keele, 1968).

As for the random generation task, a possible source of noise—suggested by the raw data—was between-subject variation in strategy. Most subjects appeared to protect randomness (H) and let number of letters generated vary between conditions, but some appeared to be doing the reverse. In Experiment 2, additional feedback and instructions were given, encouraging only randomness (H) to vary.

A second weakness in the results of Experiment 1 concerns combinations of the tone and random generation tasks. Tone task data suggested that, of the three possible concurrent tasks, random generation was the most demanding. Similarly, the random generation task data suggested that, of the three possible concurrent tasks, the tone task was the most demanding. Though consistent with a single order of general factor demand, these two observations are also consistent with the suggestion that specific interference may have occurred between the tone and random generation tasks, perhaps because of their shared auditory content. As combinations of tone and random generation tasks make a large contribution to the overall pattern of results (see Figure 2), a possible interpretation in terms of specific interference seriously weakens support for a single overall order of general factor demand. In an attempt to avoid this problem, a new tone task was devised. For Experiment 2, subjects were required to press a foot switch when they heard an infrequently occurring low-pitched tone among rapidly presented high-pitch tones. Such "oddball detection" in a rapid tone sequence is often considered to be an automatic process (Naatanen, 1985), suggesting that this task might have a reduced general factor demand. We hoped that results would still be consistent with a single order of general factor demand, but not one in which there was strong interference between tone and random generation tasks.

A third concern, supported by subjects' comments and the raw data, was that the relative demands of different tasks may have varied between subjects. For example, most subjects reported the tone task to be very demanding, but a few reported finding it easy. In Experiment 2, each subject was tested for many more days so that individual performance could be examined.

EXPERIMENT 2

Method

Subjects

Four subjects were tested; they were aged between 24 and 29, and all were university graduates. Three were female.

Design

Each subject was tested for one practice day and 12 experimental days. On the practice day, subjects were introduced to the four tasks, both singly and in combination. On each of the remaining days, 16 conditions were tested, as in Experiment 1. On any given day, the order in which the primary tasks were presented to the four subjects formed a Latin square. The order of secondary tasks was fixed for a given subject on a given day but formed a Latin square across subjects. For any given subject, the orders of primary and secondary tasks formed Latin squares over each successive block of four days.

Tasks and Procedure

There were four tasks: the manual task from Experiment 1, a modified version of random generation, a prototype learning task, and a new tone task.

Random Generation. There were a number of changes relative to that used in Experiment 1. In the instructions, randomness (H) was emphasized, and the subjects were told that the number of letters generated was not important. Secondly, feedback was given only on the randomness measure H. Furthermore, during the practice day subjects were told that the maximum score possible was 4.7, and that if their score fell below 4.0, they should try to be more random. Thirdly, on the practice day, subjects were given more practice on the task, under both single- and dual-task conditions.

Prototype Learning. In this task, subjects had to watch slight variations (exemplars) of two basic patterns (prototypes) for 3 min (learning phase) and then attempt to fill in the two underlying prototypes on two blank grids drawn on paper (test phase). Only the learning phase was done under dual-task conditions. Stimuli were presented on a computer monitor under the control of a BBC microcomputer. Each one consisted of a 4 $\times$ 4 matrix with approximately half of the 16 squares filled (light green on a dark green background). Each matrix was approximately 8 cm square and was presented in the centre of a 29 cm $\times$ 21 cm screen. The viewing distance was about 52 cm.

For each trial, two new prototypes were randomly constructed and then modified by hand to ensure that no vertical or horizontal line was either totally filled or totally empty. Prototypes always consisted of 8 filled squares. Exemplars were constructed by randomly changing 2 of the 16 squares from the corresponding prototype, such that successive exemplars were composed of 6, 8, or 10 filled squares. During the 3-min learning phase, exemplars were presented one after another in the centre of the screen. Each took approximately 500 msec to build on the screen from top left to bottom right and remained on the screen for one of three time intervals: 2 sec, 4 sec, or 6 sec. These intervals were pseudorandomly chosen so that they occurred with equal probability, allowing 40 exemplars in total to be presented in a 3-min trial. All 40 were different. Approximately half of the exemplars were generated from one prototype, and half from the other. Labels identifying the prototype pattern for each stimulus appeared under the grid in the middle; the labels used were arbitrarily assigned letters (or a number and a letter), varying from run to run.

The score for each trial was the average number of correct squares out of 16 for each of the two prototype patterns drawn. Results were scored immediately after a trial, and when the task was primary, subjects were told their score.

Tone Task. This had a conventional "oddball detection" format. Subjects were presented with a rapidly occurring sequence of tones. Only two tones were used throughout the sequence: a "target" and a background tone (approximately, 69 dB and 4700 Hz and 4900 Hz, respectively). One of these—the target—occurred infrequently. The subject's task was to respond by pressing a foot switch whenever the target was detected. However, they were told that speed of responding was not critical (again to avoid any "response-selection" bottleneck) and that any response made within 2 sec of the target would be scored as a detection.

Tones were generated by a BBC microcomputer and presented over headphones. The subjects heard 30 targets in a 3-min sequence of 240 tones. Each tone lasted 250 msec; intervals between tones were 450, 550, or 650 msec. Target occurrence was determined by a pseudorandom process. The 3-min trial was divided into 80 units, each of 2.25 sec. Within these units, each of the possible inter-tone intervals occurred once, in random order. The target could be presented only in the first tone position of any 2.25-sec unit. Given this, any response made within a 2.25-sec time unit was scored as a response to the first tone. This gave the subjects 2 sec after the offset of the first tone in which to respond. Only one (the first) response per time interval was recorded. Hits (responses when the target was presented) and correct rejections (no responses when the target was not presented) were later transformed to d' values[2] (Swets, Wilson, Tanner, & Birdsall, 1964).

Prior to any trial involving the tone task, several examples of target and background tones were played to the subject. If the tone task was primary, subjects were told their score at the end of the trial.

Instructions

The same instructions in relation to task priority were used as in Experiment 1.

[2] Zero error rates (miss or false alarm) were replaced by an estimated rate of $0.5/(N + 0.5)$, where N = number of trials (Cox & Snell, 1989).

Results

As determined by the *CIR* analysis (see further on), the estimated "best order" of tasks—in order of decreasing general factor demand—was: random generation, prototype learning, manual, tones. Thus our attempt to produce a tone task with a low estimated demand was successful. The extent to which data from each individual task were consistent with the overall "best order" may be seen by visual inspection of Figure 3, organized analogously to Figure 2. The results of the statistical analysis of this pattern of results are shown in Table 4.

Group mean performance on the random generation task is shown in the top left panel. Results are consistent with the overall "best order" in that the concurrent prototype learning task produces worse performance than the concurrent manual task, which, in turn, produces worse performance than the concurrent tone task. Single-task performance is slightly lower than the best primary performance. As expected, primary performance is better than secondary.

Group mean performance on the prototype learning task is shown in the top right panel. Results are (by visual inspection) consistent with the overall best order in that the

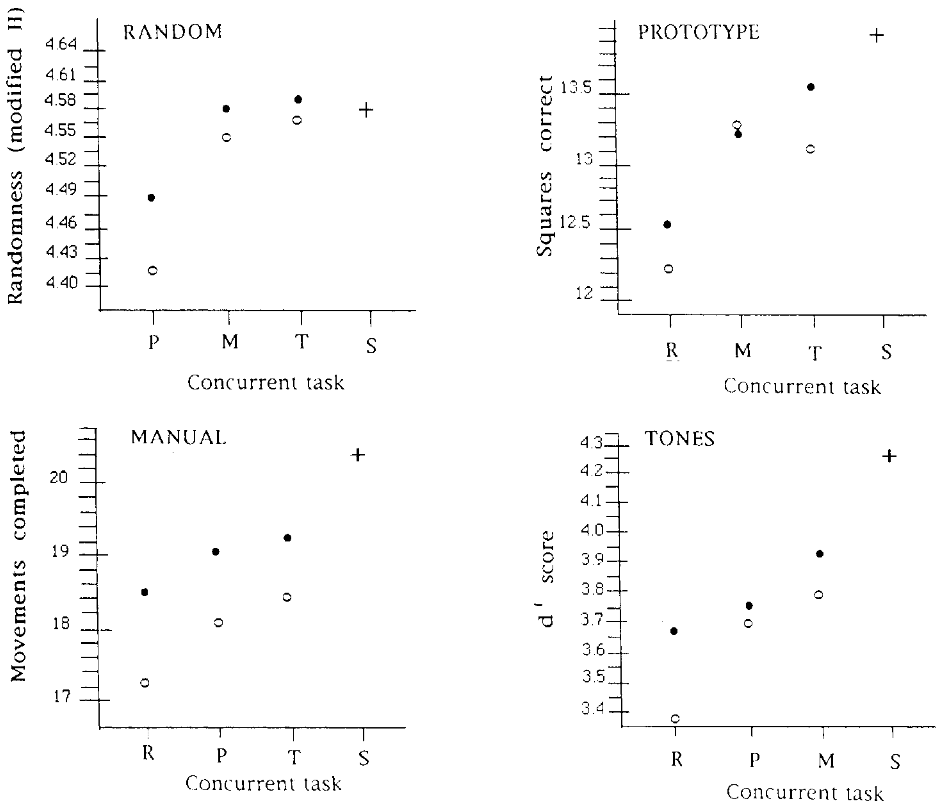

FIG. 3. Experiment 2: Average performance in all tasks and conditions. P = Prototype, R = Random, M = Manual, T = Tones, S = Single task. + = Single task score, ● = Primary task score, ○ = Secondary task score.

TABLE 4
Experiment 2. Isotonic Regressions Assessing Goodness of Fit of the Eight Orders of Decrement to a Best Order

Tasks		*SS between Tasks*	*Variance Explained*	*SO1*	*p*	*Variance Not Explained*	*S12*	*p*
Random	P	0.02	0.02	6.85	<.05	0.00	0.00	n.s.
	S	0.05	0.05	15.08	<.01	0.00	0.00	n.s.
Prototype	P	2.32	2.32	8.34	<.05	0.00	0.00	n.s.
	S	2.57	2.51	9.03	<.01	0.07	0.23	n.s.
Manual	P	1.90	1.90	10.16	<.01	0.00	0.00	n.s.
	S	2.39	2.39	12.78	<.01	0.00	0.00	n.s.
Tones	P	0.13	0.13	1.97	n.s.	0.00	0.00	n.s.
	S	0.39	0.39	5.41	<.05	0.00	0.00	n.s.

Notes: P = primary task, S = secondary task. Critical values—S01 for 3,12 observations: .10 = 2.799, .05 = 4.378, .01 = 8.950. Critical value—S12 for 3,12 observations: .10 = 3.84. S01 assesses consistency, and S12 inconsistency, with the best order.

concurrent random generation task produces worse performance than the concurrent manual task, which, in turn, produces worse performance than the tone task—at least when the prototype learning task is primary. When the prototype task is secondary, the difference between concurrent manual and tone tasks is reversed. Single-task performance is best. Primary performance is generally better than secondary, with a small reversal when the concurrent task is manual.

Group mean performance on the manual task is shown in the bottom left panel. Results are consistent with the overall best order in that the concurrent random generation task produces worse performance than the prototype learning task, which, in turn, produces worse performance than the tone task. Single-task performance is best. As expected, primary performance is better than secondary.

Group mean performance on the tone task is shown in the bottom right panel. Results are consistent with the overall best order in that the concurrent random generation task produces worse performance than the prototype learning task, which, in turn, produces worse performance than the manual task. Single-task performance is best. Primary-task performance is better than secondary.

The statistical analyses of the data shown in Figure 3, the *CIR* analysis, are shown in Table 4.[3] Of the 8 sets of means, 7 showed significant variation due to the concurrent tasks, and in all cases this variation was consistent with the single overall "best order".

For each task, an ANOVA was used to assess the effect of dual task conditions—the task when done alone, done as a primary task (averaged over the three levels), and done as a secondary task (averaged over the three levels). The effect was as follows: tones, $F(2, 6)$ =

[3] In all cases cross-validated *IR* gave exactly the same results as the overall *CIR* analysis.

62.11, $p < .0001$; random generation, $F(2, 6) = 2.96$, n.s.; prototype, $F(2, 6) = 9.81$, $p < .01$; manual, $F(2, 6) = 32.88$, $p < .0006$. Three planned contrasts were performed on each task: single compared to primary, single compared to secondary, and primary compared to secondary. The tone task was performed better as a single task than as a primary task, $F(1, 6) = 67.75$, $p < .0002$, and better as a single task than as a secondary task, $F(1, 6) = 112.86$, $p < .0001$. It was also performed almost significantly better as a primary task than as a secondary task, $F(1, 6) = 5.72$, $p < .054$. The random generation task was not performed better alone than when performed as a primary task, $F(1, 6) = 0.575$, n.s. There was almost a significant difference between when it was performed as a single task and when it was performed as a secondary task, $F(1, 6) = 5.67$, $p < .055$. There was no difference between when it was performed as a primary task and when it was performed as a secondary task, $F(1, 6) = 2.64$, n.s. The prototype task was performed better when done alone compared to when it was performed as a primary task, $F(1, 6) = 9.46$, $p < .022$, and compared to when it was performed as a secondary task, $F(1, 6) = 18.47$, $p < .005$. There was no difference between when it was performed as a primary task and when it was performed as a secondary task, $F(1, 6) = 1.49$, n.s. Performance on the manual task when done alone compared to when done as a primary task was significantly different, $F(1, 6) = 23.31$, $p < .003$. There was a difference between when it was performed alone and when it was performed as a secondary task, $F(1, 6) = 64.91$, $p < .0002$. There was also a difference when the manual task was performed as a primary task as compared to a secondary task, $F(1, 6) = 10.42$, $p < .018$.

One motivation for studying the same subjects over many days was a concern that the relative demands of different tasks might vary between subjects. To examine this, the best order for each of the four subjects was determined by individual *CIR* analyses. For two subjects, the best order was the same as the best order obtained for the group: random generation, prototype learning, manual, tones. For one subject the best order was random generation, prototype learning, tones, manual. For the fourth subject, the best order was random generation, manual, prototype learning, tones. With this subject, there were also two significant deviations from the best order. No other subject showed significant deviations from their individual best order. The variation in order between subjects, therefore, amounted to only one reversal each for two of the four subjects from the overall "best order" of the group.

Discussion

Of the 8 combinations of task with primary/secondary emphasis, 7 showed significant variation when done with different concurrent tasks. All the variation in the 7 significantly varying combinations was consistent with a single underlying order of general factor demand. Only for the tone task performed as primary was there no significant variation between conditions. Our attempted improvements over Experiment 1 were successful: The prototype learning task was sensitive to concurrent demands, and the demands of the tone task were low rather than high. (Of course, as a consequence, the tone task was itself relatively insensitive to the secondary task.) Finally, in almost all cases, single-task performance was best, primary performance intermediate, and secondary performance worst.

GENERAL DISCUSSION

The results we have described provide strong support for the idea that some "general factor" is involved in dual-task performance decrement. The results are unlikely to have been obtained had there been only large specific interference effects of the kind described in the introduction. In order for an interpretation in terms of such specific interference to be valid, one would have to postulate that a collection of specific interference effects had emerged in both these experiments that by chance mimicked the predicted pattern.

The conventional types of strong specific interference do not seem a plausible interpretation of the current results. However, the current results and ideas along the lines of specific interference can be reconciled within a new framework. Assume that there are sources of specific interference that are small but very numerous and very widespread. As such, these would be very unlike the traditional notion of specific interference effects discussed in the introduction (e.g. common input modality, Treisman & Davies, 1973). Such specific interference effects have typically been supposed to be few, producing very large dual-task interference with a fairly obvious source, such as having a common input modality.

This contrasts with an explanation along the lines of "multiple, widespread, but small" specific interference effects, such as that which can be drawn from Navon (1984). He argues that doing a task may have so many specific side-effects that it will interfere with very many different tasks. The more complex a task, the greater will be the number of side-effects, and so the greater will be the dual-task interference. There may be different specific interferences in different task combinations, but a more complex task should always interfere more than a less complex task with any concurrent task. As long as certain strong sources of interference are avoided (e.g. a common response selection bottleneck, Pashler, 1990), the reason for dual-task interference will not be competition for a single important cognitive process or stage of processing. Rather, dual-task interference will be related more to the *quantity* of cognitive processes that are involved. Such a distributed interference model may offer a plausible theoretical interpretation of the general factor. As the data require, each task is associated with a unique quantity—the quantity of cognitive processes or side-effects—determining how strongly it interferes with *any* dissimilar concurrent task. Such an account can be placed alongside "general resource" and "central executive" interpretations of general dual-task interference.

The results were also largely consistent with the weighting model for general factor allocation. As expected, primary-task performance was not perfectly protected, but was usually better than secondary-task performance. Critically, the weighting model led us to predict that the same order of interference would be seen in primary- and secondary-task data, as, indeed, we found. The data suggest that the weighting model is at least approximately correct—each task has an intrinsic weight and an additional small weight introduced by instruction. This weight introduced by instruction appears to remain roughly constant, irrespective of the task with which it is combined. Had it not remained so, then we might have expected a violation of the consistent order.

One of the interesting results of this study is that novel large specific sources of interference of the conventional kind were apparently avoided. If any had existed, the general factor pattern would have been masked. That no new sources of strong specific interference were found suggests that the major sources may already have been identified in previous studies and thus can be avoided.

The nature of the general factor is not identified by this study. In addition to the cognitive complexity account considered above, the results are consistent with the various general factor models discussed in the introduction. The general factor may be a limited pool of processing resource that needs to be invested for a task to be performed. It may be a limited central executive that coordinates or monitors other processes and is limited in how much it can deal with at one time. It may also represent a general limit of the entire cognitive system on the amount of information that can be processed at a given time. The method developed here deals only with the existence of a general factor in dual-task decrement, not its nature.

For 20 years the debate has continued over whether dual-task interference is best explained by "general factor" or "specific interference" models (e.g. Allport, 1980; Allport, Antonis, & Reynolds, 1972; McLeod & Posner, 1984). The present results suggest that the disagreement is misplaced. Though evidence for specific conflicts is overwhelming, a further general source of interference is revealed when these are excluded, suggesting that both general and specific limits can apply.

REFERENCES

Ackerman, P.L., Schneider, W., & Wickens, C.D. (1984). Deciding the existence of a time-sharing ability: A combined methodological and theoretical approach. *Human Factors*, *26*, 71–82.

Allport, D.A. (1980). Attention and performance. In G. Claxton (Ed.), *Cognitive psychology: New directions* (pp. 112–153). London: Routledge and Kegan Paul.

Allport, D.A., Antonis, B., & Reynolds, P. (1972). On the division of attention: A disproof of the single-channel hypothesis. *Quarterly Journal of Experimental Psychology*, *24*, 225–235.

Baddeley, A.D. (1966). The capacity for generating information by randomization. *Quarterly Journal of Experimental Psychology*, *18*, 119–129.

Baddeley, A.D., & Lieberman, K. (1980). Spatial working memory. In R.S. Nickerson (Ed.), *Attention and performance VIII* (pp. 521–539). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Baddeley, A.D., & Hitch, G. (1974). Working memory. In C. Bower (Ed.), *Recent advances in learning and motivation, Vol VII* (pp. 45–90). London: Academic Press.

Broadbent, D.E. (1958). *Perception and communication*. London: Pergamon Press.

Broadbent, D.E. (1982). Task combination and selective intake of information. *Acta Psychologia*, *50*, 253–290.

Colavita, F., & Weisberg, D. (1979). A further investigation of visual dominance. *Perception and Psychophysics*, *25*, 345–347.

Cox, D.R., & Snell, E.J. (1989). *Analysis of binary data* (2nd ed.) London: Chapman & Hall.

Duncan, J., Williams, P., Nimmo-Smith, I., & Brown, I. (1993). The control of skilled behavior: Learning, intelligence and distraction. In D. Meyer and S. Kornblum (Eds.), *Attention and performance XIV*. New York: MIT Press.

Eriksen, C.W., & Johnson, H.J. (1964). Storage and decay characteristics of non-attended auditory stimuli. *Journal of Experimental Psychology*, *68*, 37–43.

Hirst, W., & Kalmar, D. (1987). Characterizing attentional resources. *Journal of Experimental Psychology: General*, *116*, 68–81.

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

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

McLeod, P., & Posner, M.I. (1984). Privileged loops from percept to act. In H. Bouma and D.G.

Bouwhuis (Eds.), *Attention and Performance X* (pp. 55–66). Hove, U.K.: Lawrence Erlbaum Associates Ltd.

Moray, N. (1967). Where is capacity limited? A survey and a model. *Acta Psychologia*, *27*, 84–92.

Murdock, B.B. (1965). Effects of a subsidiary task on short-term memory. *British Journal of Psychology*, *56*, 413–419.

Naatanen (1985). Stimulus processing: In event-related potentials magnetoencephalogram and regional blood flow. In M. Posner & O. Martin (Eds.), *Attention and Performance XI*. Hove, U.K.: Lawrence Erlbaum Associates Ltd.

Navon, D. (1984). Resources—a theoretical soup store? *Psychological Review*, *91*, 216–234.

Navon, D., & Gopher, D. (1979). On the economy of the human-processing system. *Psychological Review*, *86*, 214–255.

Neisser, U., & Becklen, R. (1975). Selective looking: Attending to visually specified events. *Cognitive Psychology*, *7*, 480–494.

Norman, D.A., & Bobrow, D.J. (1975). On data-limited and resource-limited processes. *Cognitive Psychology*, *7*, 44–46.

Norman, D.A., & Shallice, T. (1980). *Attention to action: Willed and automatic control of behaviour* (Chip 99, report no. 8006). San Diego, CA: University of California, Center for Human Information Processing.

Pashler, H. (1990). Do response modality effects support multiprocessor models of divided attention? *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 826–842.

Posner, M.I. (1978). *Chromometric explorations of mind*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Posner, M.I., & Keele, S.W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, *77*, 353–363.

Posner, M.I., Sandson, J., Dhawan, M., & Shulman, G.L. (1989). Is word recognition automatic? A cognitive anatomical approach. *Journal of Cognitive Neuroscience*, *1*, 50–60.

Reisberg, D. (1983). General mental resources and perceptual judgments. *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 966–979.

Robertson, T., Wright, F.T., & Dykstra, R.L. (1988). *Order-restricted statistical inference*. Chichester, NY: John Wiley.

Salame, P., & Baddeley, A.D. (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.

Shallice, T., McLeod, P., & Lewis, K. (1985). Isolating cognitive modules with the dual-task paradigm: Are speech perception and production separate processes? *The Quarterly Journal of Experimental Psychology*, *37*, 507–532.

Swets, J.A., Wilson, P., Tanner, Jr., & Birdsall, T.C. (1964). Decision processes in perception. In J.A. Swets (Ed.), *Signal detection and recognition by human observers: Contemporary readings*. New York: Wiley.

Treisman, A.M., & Davies, A. (1973). Divided attention to ear and eye. In S. Kornbum (Ed.), *Attention and performance IV* (pp. 101–117). London: Academic Press.

Tune, G.S. (1964). Sequential errors in a time-sharing task. *British Journal of Psychology*, *54*, 415–419.

Yee, P.L., Hunt, E., & Pellegrino, J.W. (1991). Coordinating cognitive information: Task effects and individual differences in integrating information from several sources. *Cognitive Psychology*, *23*, 615–680.

Original manuscript submitted 2 February 1994
Accepted revision received 10 July 1995