

# The cost of expecting events in the wrong sensory modality

CHARLES SPENCE

*University of Oxford, Oxford, England*

MICHAEL E. R. NICHOLLS

*University of Melbourne, Melbourne, Australia*

and

JON DRIVER

*University College London, London, England*

We examined the effects of modality expectancy on human performance. Participants judged azimuth (left vs. right location) for an unpredictable sequence of auditory, visual, and tactile targets. In some blocks, equal numbers of targets were presented in each modality. In others, the majority (75%) of the targets were presented in just one expected modality. Reaction times (RTs) for targets in an unexpected modality were slower than when that modality was expected or when no expectancy applied. RT costs associated with shifting attention from the tactile modality were greater than those for shifts from either audition or vision. Any RT benefits for the most likely modality were due to priming from an event in the same modality on the previous trial, not to the expectancy per se. These results show that stimulus-driven and expectancy-driven effects must be distinguished in studies of attending to different sensory modalities.

Can people selectively direct their covert attention to a particular sensory modality, to the exclusion of other modalities? There has been renewed interest in this classic question, as researchers examine possible deficiencies in shifting attention between modalities for various clinical disorders, including schizophrenia, autism, and Huntington's disease (e.g., Ciesielski, Knight, Prince, Harris, & Handmaker, 1995; Ferstl, Hanewinkel, & Krag, 1994; Sprengelmeyer, Lange, & Hömberg, 1995), plus the neural consequences of attending to one modality versus another (e.g., see Driver & Spence, 1998; Foxe, Simpson, & Ahlfors, 1998; Steinmetz et al., 2000). Although seminal studies apparently showed that attended modalities are preferentially processed (e.g., Klein, 1977; Posner, Nissen, & Klein, 1976), many past results remain open to nonattentional explanations in terms of criterion shifting, repetition priming, or spatial confounds. We briefly describe these potential confounds and then report a new experiment designed to overcome them, via a modified version of Klein's method.

Simple *detection* latencies for target stimuli can be faster when participants are correctly cued to the target modality in advance than when they divide their attention between several modalities or when they are cued to another modality (e.g., Boulter, 1977; Post & Chapman, 1991). However, this may reflect a criterion shift for stimuli in the expected modality, rather than an increased sensitivity (cf. Duncan, 1980; Spence & Driver, 1997). Some modality-cuing effects have been observed in *discrimination* tasks (e.g., Gescheider, Sager, & Ruffolo, 1975; Klein, 1977; Posner et al., 1976), allowing the assessment of any criterion shifts. For example, participants in Klein's study made speeded left/right responses for auditory, visual, and tactile-kinesthetic targets. Attention was manipulated by presenting the majority of targets for each block in just one modality. Discrimination latencies were significantly faster for targets in the currently expected target modality; accuracy was also better, making a criterion shift implausible.

However, Spence and Driver (1997) highlighted another potential problem with most previous modality-cuing studies (e.g., Boulter, 1977; Gescheider et al., 1975; Klein, 1977; Posner et al., 1976; Post & Chapman, 1991; Roland, 1982): Targets in the various modalities were presented from different spatial locations. In Klein's study, auditory targets were presented over headphones, visual targets from a cathode-ray tube in front of participants, and tactile-kinesthetic stimuli to their right forefinger (which was moved to the left or right by a servo motor). Consequently, information about the likely target

---

This work was supported by a Programme grant from the Medical Research Council (U.K.) and by a grant from the Australian Research Council. Our thanks to Nicole Gillespie for running the participants in the experiment. We are also deeply indebted to Ray Klein for suggesting many of the data analyses included in this manuscript and for his encouragement throughout the review process. Correspondence should be addressed to C. Spence, Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, England (e-mail: charles.spence@psy.ox.ac.uk).

modality may have directed spatial attention (overtly and/or covertly) to the likely *location*, rather than to one modality per se (cf. Posner, Nissen, & Ogden, 1978, p. 152; Spence & Driver, 1997). In this study, we repeat Klein's basic design, while correcting for the potential location confound. We also adapted the design and analyses in order to address several further issues.

In particular, we wished to determine any asymmetries between the various modalities (vision, hearing, and touch) in the effects of expecting or not expecting events within them. Careful examination of this seemed important, given recent claims that clinical populations may show altered patterns of sensory dominance in this respect (e.g., Zahn, Pickar, & Haier, 1994). Several influential results within the normal *visual dominance* literature (e.g., Klein, 1977; Posner et al., 1976) suggest that asymmetries may exist between attending to different modalities, but the interpretation of these past results remains open to question. For example, Posner et al. (1976) reported that modality-cuing effects were smaller for visual than for tactile targets. This could have several explanations. It might be due, as Posner et al. (1976) concluded, to a default tendency to attend to vision. Or, as the same authors also considered (p. 164), it could be due instead to some peculiar difficulty in shifting attention *away* from vision, not to some preference for directing attention *toward* vision (cf. LaBerge, 1973). Finally, it seems possible that the result might simply be an artifact of the manner in which the likely modality was indicated. As in many other studies (e.g., Boulter, 1977; Gescheider et al., 1975; Post & Chapman, 1991; Spence & Driver, 1997), the likely target modality was signaled for each trial by a *visual* cue (moreover, feedback at the end of the trial was also presented visually). This alone might have biased attention toward the visual modality in these studies. We examined whether so-called *visual dominance* effects still occur when no visual cues are used. To this end, we adapted Klein's blocked-cuing paradigm, in which participants are instructed how to allocate their attention for an entire block of trials, not on a trial-by-trial basis. Unlike Posner et al. (1978), but like Klein, we used three possible modalities rather than two, since only this can determine whether the pattern of results for an unexpected modality is specific to the modality that attention must be shifted *from* or to the modality it must be shifted *toward*. Unlike Klein, we also included a *divided-attention* condition, intended as a neutral baseline for calculating whether modality-based attention can produce benefits for the expected modality, as well as costs for the unexpected. Finally, as will be explained subsequently, some of our analyses took into account the nature of the preceding trial, in an effort to separate expectancy effects from those that are due merely to passive exogenous priming (Epstein & Rock, 1960) by the modality of the preceding target. These have been confounded in several previous studies of attending to one sensory modality versus another.

## METHOD

### Participants

Twenty participants (mean age of 22 years) were recruited through advertisement for undergraduate credit.

### Apparatus and Materials

The participants sat at a table in a dark room with arms resting on the tabletop. Head movements were precluded by an adjustable chinrest. A green light-emitting diode (LED) served as a fixation point, placed directly in front of the participants' eyes (note that, in any case, saccades should not influence any attentional effects observed, because target location remained equally unpredictable regardless of expected modality). Two circular loudspeakers (12.7-cm diameter) could each present auditory targets, one placed on either side of fixation at an eccentricity of 41°. Although this eccentricity is larger than those conventionally used in studies of purely visual attention, it is within the range used in past studies of crossmodal attention (see Spence & Driver, 1997), and its use here facilitated auditory localization for the target task. The fixation light and target loudspeakers were placed on a virtual circle (radius = 50 cm) at eye level centered on the participant's head. One red circular LED (with a diameter of 1 cm and a luminance of 64 cd/m<sup>2</sup>) was placed immediately in front of each target loudspeaker to present visual targets, and a tactile stimulator was placed directly behind each loudspeaker to present tactile targets. This arrangement ensured that all the auditory, visual, and tactile targets were presented from the same two possible positions, thus removing the confound between modality and location that has plagued previous studies.

The participants rested their left index fingers over a circular aperture in the tactile stimulator behind the left loudspeaker and their right index fingers over a stimulator behind the right loudspeaker. The participants were unable to see their hands, but velcro strips ensured a constant contact between finger and stimulator. Tactile targets consisted of the presentation of a circular rod (0.7 mm in diameter) to the underside of one index finger, by raising it through the aperture in the tactile stimulator (travel time for the rod was 5 msec; the moving rod was capable of fully displacing a weight of up to 5 g). Auditory targets were white noise presented at 90 dB(A) from one loudspeaker, and visual targets were the illumination of one LED. All the targets were presented for 50 msec. White noise was presented continuously throughout the experiment from a loudspeaker situated under the participants' chair at 70 dB(A), to mask the slight noises caused by making a response or by the operation of the tactile stimulators. The participant depressed two footpedals with the right foot (one under the toes and the other under the heel) throughout each experimental block and responded by briefly releasing one or the other pedal (lifting their toes for right targets and their heels for left targets). This stimulus-response mapping was used because the hands were occupied with the tactile stimulators; the pedal responses have been successfully used in our previous crossmodal studies (see Spence, Nicholls, Gillespie, & Driver, 1998).

### Design

The two within-subjects factors were target modality (three levels; auditory, visual, or tactile, all intermingled in a random sequence) and block type (four levels: expect audition, expect vision, expect touch, or divided attention). In each expect audition block, there were 36 (75%) auditory targets, 6 (12.5%) visual targets, and 6 (12.5%) tactile targets, all in a random sequence. Comparable target probabilities were used in the expect vision (i.e., 75% visual, 12.5% auditory, and 12.5% tactile) and expect touch (75% tactile, 12.5% visual, and 12.5% auditory targets) blocks, whereas equal numbers of auditory, visual, and tactile targets were presented in the divided-attention blocks of trials. There were 4 blocks of 20

practice trials (which were not analyzed), followed by 16 test blocks of 48 test trials each.

### Procedure

The participants were informed about the probability of occurrence of each type of target for the upcoming block. They were instructed to divide their attention equally between the three modalities at the beginning of the divided-attention blocks, and to attend principally to the most likely target modality in the other blocks (i.e., they still had to *respond* to targets presented in the other less common modalities but should concentrate on the most likely modality to optimize performance for those targets). The participants were instructed to make the left/right discrimination response as rapidly and accurately as possible to target events in every modality, and the trial was terminated if no response had been made within 1,500 msec of target onset. The interval between successive targets was randomly in the range of 1,500–1,900 msec. Error feedback was presented if the participant's response was incorrect, or if no response had been made by the time the trial terminated. This feedback consisted of the pulsed simultaneous presentation of *all* six possible target signals (i.e., two auditory, two visual, and two tactile) in six 10-msec bursts, each separated by 30-msec intervals, to ensure stimulation of all three modalities. Targets were equally likely on the left and the right for each modality in each block. All four block types were presented once in each successive quadruplet of blocks, with their order of presentation randomized within each quadruplet (with the constraint that the same block type was never presented on successive blocks across quadruplets).

## RESULTS AND DISCUSSION

Trials with an incorrect response and those immediately succeeding an incorrect response were discarded from the reaction time (RT) analyses. The mean RTs (and associated error percentages) are shown for each condition in Table 1, together with any benefits of expecting the target modality or any costs owing to expecting a dif-

ferent modality. As Table 1 shows, both costs and benefits were present for every case.

A within-subjects analysis of variance (ANOVA), with factors of block type (4)  $\times$  target modality (3), on the RT data revealed a main effect of target modality [ $F(2,38) = 4.3, p = .02$ ], with slower responses to tactile targets overall ( $M = 520$  msec) than to visual ( $M = 513$  msec) or auditory targets ( $M = 504$  msec). There was a main effect of block type [ $F(3,57) = 11.2, p < .0001$ ], with faster responses overall in the divided-attention blocks, owing mainly to costs for the two unexpected modalities that contributed two thirds of the means in the other blocks. The interaction between block type and target modality was also significant [ $F(6, 114) = 45.7, p < .0001$ ], owing to costs for the unexpected modalities and benefits for the expected, relative to the divided-attention baseline. The significance of paired *t* tests on each cost or benefit is given in the note to Table 1. Note also that the largest costs of expecting the 'wrong' modality all involve touch (i.e., shifting attention away from touch or toward touch seems to carry the greatest penalty, as will be discussed later). A similar analysis of the error data revealed no main effects or interactions (all  $F$ s  $< 1$ ).

These results are similar to those of Klein's (1977, Experiment 1) study (see the first and second columns of numbers in Table 2), whose design we had adapted in order to exclude the spatial confound and to add the divided-attention baseline. This similarity is quite striking, given the different stimuli used. Since Klein did not have a neutral baseline condition, he could only measure costs-plus-benefits (i.e., the difference in performance for a given modality when expected vs. unexpected). His results for this measure all fall within the 95% confidence interval of our own unexpected-minus-expected findings, except for the particular cases of expecting vision but getting audition and vice versa, where Klein reports modality-cuing effects nearly twice those in the present experiment. This difference probably reflects the spatial confound inherent in Klein's study (i.e., audition over headphones, vision from in front of the participant; see Spence & Driver, 1997, for further evidence that such confounds can artifactually elevate modality-cuing effects between hearing and vision).

### Endogenous Versus Exogenous Components of Attention to a Modality

Spence and Driver (1997) recently noted that standard analyses for studies with blocked modality-probabilities, which include all trials (as for the analyses presented above and those of Klein, 1977, and Posner et al., 1976), may fail to distinguish the critical effects of *expectancy* from *stimulus-driven* effects, such as intertrial repetition effects (see also Epstein & Rock, 1960, on this distinction). Past studies (e.g., Ferstl et al., 1994; Zubin, 1975) have shown that people typically respond more slowly to a target in a given modality if the preceding target occurred in a different modality (thus producing a *cross-modal* trial), rather than in the same modality (resulting

**Table 1**  
Mean Reaction Times (RTs; in Milliseconds), Their Standard Errors (in Parentheses), and Percentages of Errors for Auditory, Visual, and Tactile Targets on all Trials (i.e., Including Both Ipsimodal and Crossmodal Trials) as a Function of Block Type, Showing the Attention Effects

Target Modality		Block Type			
		Expect Auditory	Expect Visual	Expect Tactile	Divided Attention
Auditory	RT	454 (15)	522 (18)	552 (22)	487 (18)
	%	2.8	3.1	4.0	3.1
<i>Attention effect</i>		<b>+34†</b>	-35*	-65†	
Visual	RT	522 (17)	467 (14)	571 (24)	492 (17)
	%	3.6	3.8	4.8	3.7
<i>Attention effect</i>		-31†	<b>+25*</b>	-80†	
Tactile	RT	549 (19)	548 (22)	482 (18)	499 (20)
	%	3.7	4.7	3.3	3.4
<i>Attention effect</i>		-50†	-49*	<b>+17*</b>	

Note—The benefits for an expected modality were calculated by subtracting the RT for a particular target when expected from that for the same target in the divided-attention blocks (the results for this appear in the main top-left/lower-right diagonal as boldface figures). The costs for an unexpected modality were calculated by subtracting the RT for a particular target when a different modality was expected from that for the divided-attention blocks (italic figures). \* $p < .01$ , † $p < .001$ , by *t* test pairwise comparison.

**Table 2**  
**Summary of Attention-Switching Effects for an Expected Modality Versus the Same Modality When Unexpected (Costs Plus Benefits Combined; in Milliseconds) as Reported by Klein (1977) and as Found in the Present Study**

Shifting Attention From	Shifting Attention To	Klein (1977)	Experiment 1	
		Ipsimodal + Crossmodal	Ipsimodal + Crossmodal	Crossmodal Trials Only
vision	audition	106	68	37
touch	audition	104	98	69
audition	vision	102	55	39
touch	vision	109	104	86
vision	touch	50	66	36
audition	touch	63	67	42

Note—The elimination of ipsimodal trials in our analysis led to a reduction in attention-switching effects for all conditions.

in an *ipsimodal* trial). This purely stimulus-driven effect has become known as the modality-shifting effect (MSE). In the present experiment, the most likely modality within a block of trials will also tend to be involved in more ipsimodal than crossmodal trials (cf. Kornblum, 1967), whereas the reverse will apply to the less likely modalities. The faster RTs for targets in the expected (vs. unexpected) modality within the present experiment might therefore be due merely to a stimulus-driven MSE (i.e., an exogenous benefit when the same modality is repeated across successive trials), not to an endogenous expectancy. This problem also applies in principle to the results of Klein's (1977) study, and analysis of our own data highlights the reality of the problem.

The mean RTs (and associated error percentages) for auditory, visual, and tactile targets as a function of the modality of the target on the preceding trial are shown in Table 3. To assess stimulus-driven MSEs in our own study, a within-subjects ANOVA (target modality (3) × preceding target modality (3) × block type [2; focused vs. divided attention]) was conducted on the data summarized in Table 3. The analyses of both RT and error data revealed main effects of the preceding target modality [ $F(1,19) = 4.9, p = .01$ , for RT data;  $F(1,19) = 4.3, p = .02$ , for the error data], caused by participants' responding more slowly and less accurately following a tactile target on the preceding trial ( $M = 490$  msec and 4.0% errors) than following either auditory (483 msec and 2.9% errors) or visual ( $M = 483$  msec and 3.1% errors) targets. The interaction between the preceding target modality and the current target modality was also significant in the RT data [ $F(1,19) = 35.8, p < .001$ ]. Paired-samples  $t$  tests revealed that performance on ipsimodal trials was significantly faster than that on crossmodal trials for all possible comparisons ( $p < .01$ ) across all three target modalities, demonstrating that exogenous MSEs can significantly influence performance in the present left/right spatial discrimination task. None of the other terms in the analysis of either the RT or the error data reached significance.

The results of these analyses support Spence and Driver's (1997) suggestion that the supposed effects of modal-

ity expectancy in Klein (1977) and Posner et al. (1976) might be contaminated by exogenous MSEs. (Unfortunately, the raw data from Klein's study are no longer available for further analysis; Klein, personal communication, August 5, 1999.) Moreover, the present results demonstrate tactile MSEs for the first time (all previous studies have focused solely on auditory and visual stimuli). In fact, MSEs for auditory targets following a tactile target on the preceding trial were significantly larger than those elicited by a preceding visual target (mean RTs of 505 and 479 msec, respectively;  $p < .01$  by  $t$  test pairwise comparison). There was a similar trend for larger MSEs following tactile targets (mean RT of 496 msec) than following auditory targets (mean RT of 489 msec) in the visual target data too, although this comparison failed to reach significance. For tactile targets, there was absolutely no difference between the MSEs elicited by auditory or visual preceding target stimuli (mean RTs of 511 and 512 msec, respectively; n.s.). It appears, therefore, that stimulus-driven MSEs may be greater following tactile stimuli than following either auditory or visual stimuli.

We conducted a further analysis of the data to assess (and reduce) the influence of such exogenous factors on our expectancy-based modality-cuing effects. In this new analysis, we simply eliminated all ipsimodal trials (same modality as the previous trial), to prevent these from advantaging the expected modality, owing to its being more likely to contain ipsimodal trials. Any residual cuing effects from modality expectancy within just the *crossmodal* trials should provide a purer measure of *endogenous* attention to a modality, uncontaminated by exogenous MSEs from the preceding trial.<sup>1</sup> To provide an appropri-

**Table 3**  
**Mean Reaction Times (RTs; in Milliseconds), Their Standard Errors (in Parentheses), and Percentages of Errors, for Auditory, Visual, and Tactile Targets as a Function of Block Type and Modality of the Preceding Target**

Target Modality	Preceding Target Modality		Block Type			
			Expect Auditory	Expect Visual	Expect Tactile	Divided Attention
Auditory	auditory	RT	438 (14)	—	—	459 (19)
		%	2.9	—	—	3.3
	visual	RT	469 (16)	—	—	488 (19)
		%	1.7	—	—	2.0
	tactile	RT	505 (20)	—	—	505 (18)
		%	3.2	—	—	3.9
Visual	auditory	RT	—	481 (15)	—	498 (18)
		%	—	3.1	—	2.8
	visual	RT	—	454 (15)	—	464 (17)
		%	—	3.7	—	4.5
	tactile	RT	—	495 (16)	—	498 (19)
		%	—	5.0	—	4.1
Tactile	auditory	RT	—	—	509 (21)	513 (23)
		%	—	—	2.8	2.7
	visual	RT	—	—	519 (23)	502 (21)
		%	—	—	3.5	3.0
	tactile	RT	—	—	466 (17)	473 (19)
		%	—	—	3.4	4.7

Note—Cells where insufficient trials were available for analysis are indicated by dashes.

**Table 4**  
**Mean Reaction Times (RTs; in Milliseconds),**  
**Their Standard Errors (in Parentheses), and Percentages**  
**of Errors for Auditory, Visual, and Tactile Targets on**  
**Just the Crossmodal Trials, as a Function of Block Type,**  
**Showing the Attention Effects**

Target Modality		Block Type			
		Expect Auditory	Expect Visual	Expect Tactile	Divided Attention
Auditory	RT	487 (18)	524 (18)	556 (22)	497 (19)
	%	2.4	3.2	3.9	3.0
<i>Attention effect</i>		<b>+10</b>	-27*	-59‡	
Visual	RT	526 (18)	488 (15)	573 (25)	497 (19)
	%	3.1	4.0	4.8	3.5
<i>Attention effect</i>		-29‡	<b>+10</b>	-76‡	
Tactile	RT	556 (20)	549 (22)	513 (21)	507 (22)
	%	3.6	4.5	3.0	2.8
<i>Attention effect</i>		-48‡	-42‡	<b>-6</b>	

Note—The benefits for an expected modality were calculated by subtracting the RT for a particular target when expected, from that for the same target in the divided-attention blocks (the results for this appear in the main top-left/lower-right diagonal, as boldface figures). The costs for an unexpected modality were calculated by subtracting the RT for a particular target when a different modality was expected, from that for the divided-attention blocks (italic figures). \* $p < .05$ , † $p < .01$ , ‡ $p < .001$ , by  $t$  test pairwise comparison.

ate baseline for the subsequent analyses, we similarly eliminated ipsimodal trials from the divided-attention blocks (which were unique in having more crossmodal than ipsimodal trials for every modality).

The mean RTs for just the crossmodal trials, together with the corresponding error rates, are shown in Table 4 (and summarized in the third column of numbers in Table 2), with the size and significance of the various attention effects indicated. An ANOVA (block type [4]  $\times$  target modality [3]) on the RT data for crossmodal trials again revealed a main effect of target modality [ $F(2,38) = 3.5, p = .04$ ] caused by slower responses to tactile targets overall ( $M = 531$  msec) than to visual ( $M = 521$  msec) or auditory ( $M = 516$  msec) targets. There was again an effect of block type [ $F(3,57) = 10.3, p < .0001$ ], with faster responses overall in the divided-attention blocks, owing to costs for the two unexpected modalities in the other blocks. Finally, the interaction between block type and target modality was again highly significant [ $F(6,114) = 23.4, p < .0001$ ], owing to the various attention effects discussed below. A comparable ANOVA on the error data found no significant terms.

Note that RTs were similar for all modalities in the divided-attention condition (see the rightmost column in Table 3 and also in Table 1). Pairwise comparisons found no difference in speed or accuracy between any of the modalities for these blocks. This shows that the participants made the left/right discrimination equally rapidly for targets in each modality, so that the stimuli were well matched in this respect. Responses were significantly faster for auditory targets than for visual or tactile targets in the expect auditory blocks ( $p < .01$ , for both  $t$  test pairwise comparisons). Similar effects (all at  $p < .01$ ) were

found concerning visual targets in the expect vision blocks and tactile targets in the expect touch blocks. It is this pattern which produces the interaction between target modality and block type in the ANOVA above.

The various attention effects are shown in boldface and/or italics in Table 4. As in the analysis for Table 1 (which differed in that it included ipsimodal trials), significant costs of attending to the wrong modality were found in every case. Moreover, the participants also tended to make more errors overall when their attention had been directed to the wrong modality, ruling out simple criterion-shifting explanations for the RT costs.

Unlike the previous analyses in Table 1, no reliable *benefits* were now found for targets in the expected modality, as compared with divided attention (see main top-left/lower-right diagonal of figures in boldface in Table 4).<sup>2</sup> This change in the pattern of results suggests that the apparent benefits of attending to the “correct” modality seen in Table 1 may be due to the stimulus-driven MSE effect, caused by intertrial repetition of the same stimulus modality (which will inevitably benefit the most probable modality in any block most often). It thus appears that while stimulus repetition can produce benefits in performance, expectancy may only produce costs for the unexpected modality, not significant benefits for the expected modality. This aspect of our results accords with Pashler’s (1998; see pp. 190–191) account of the effects of attentional set (see also Dai, Scharf, & Buus, 1991, Gawryszewski, Riggio, Rizzolatti, & Umiltà, 1987, Shiu & Pashler, 1994, Spence & Driver, 1996, and Spence, Pavani, & Driver, 2000, for other examples of expectancy costs in the absence of expectancy benefits).

As in Table 1, the data in Table 4 suggest that the largest costs are associated with a violated *tactile* expectation. Indeed, RTs for auditory targets were significantly slower [ $F(1,19) = 8.2, p = .01$ ] when a tactile target was expected ( $M = 556$  msec, 3.9% errors) than when a visual target was expected ( $M = 524$  msec, 3.2%), suggesting a larger cost of shifting attention away from touch. Similarly, RTs to visual targets were also significantly longer [ $F(1,19) = 11.7, p = .003$ ] when a touch was expected ( $M = 573$  msec, 4.8%) than when a sound was expected ( $M = 526$  msec, 3.1%). This suggests that people find it more difficult to shift attention *away* from the tactile modality than from either audition or vision. Given the equivalent performance on our task for the three modalities in the divided-attention condition, this seems unlikely to reflect any difficulty peculiar to the present tactile task. It may, therefore, reflect a general asymmetry in attending to the modalities, with people finding it particularly hard to shift attention away from touch.

There was no significant difference [ $F(1,19) = 0.3, n.s.$ ] between RTs for tactile targets when expecting auditory ( $M = 556$  msec, 3.6%) versus visual ( $M = 549$  msec, 4.5%) targets, suggesting that the costs of shifting attention *from* audition or vision are similar to each other (as was also reported by Klein, 1977, and Spence & Driver, 1997). Finally, although visual inspection of Table 4 sug-

gests that the costs associated with shifting attention *toward touch* from either audition or vision (mean costs of  $-48$  and  $-42$  msec, respectively) may be larger than those associated with shifting attention to *audition* from vision or to *vision* from audition (mean costs of  $-27$  and  $-29$  msec, respectively), these differences failed to reach statistical significance.

Our finding that expectancies for touch have greater effects than those for vision replicates Posner et al. (1978). The novel result is that tactile expectancies also have larger influences than do those for audition. This suggests that people specifically find it hard to shift attention away from the tactile modality, rather than simply having a general predisposition to attend toward vision, as Posner et al. (1976) had previously suggested. Their suggestion was based on an experiment with only two modalities. The present use of three modalities within the same study (like Klein, 1997) allows a fuller picture and, moreover, goes beyond Klein's pioneering study by (1) not confounding modality and location in space, (2) including a baseline condition (the divided-attention blocks), where each modality was equally likely, and (3) analyzing for stimulus-driven effects owing to immediate repetition of a modality.

The present conclusion that shifting attention from the tactile modality (both exogenously and endogenously) is particularly difficult may relate to Roland's (1982, p. 1068) PET finding that shifts of attention away from the tactile modality produce qualitatively different changes in the pattern of regional cerebral blood flow than those seen when shifting attention away from audition or vision (the latter reduced activity in appropriate unimodal cortices, whereas the former did not). Moreover, Eimer and Driver (2000) recently found with event-related potentials (ERPs) that modality selection behaves uniquely for touch, in that only touch can be spatially decoupled from spatial attention in the other two modalities when it is task irrelevant.

It is interesting to speculate about the possible reason for the present difficulty in shifting attention away from the tactile modality. One possibility, previously raised by philosophers (see Martin, 1995), is that an "internal" rather than an "external" perspective may be adopted with touch, whereas sounds and sights must always be experienced as properties of the external world. That is, one can experience a tactile sensation as existing on the body's surface in a way that is not possible for visual stimulation on the retina or auditory stimulation on the cochlea. If this is indeed the case, the extra cost of shifting attention from the tactile modality might conceivably reflect the difficulty of shifting attention from a representation of the body (what Stein, 1989, calls *personal space*) to the peripersonal or extrapersonal representations in which auditory and visual events are coded (see Martin, 1995). In this regard, it may be revealing to examine modality attention for cases involving *active* touch (rather than

passive touch, as used here), when the hands are used to explore the external world, so that tactile attention may then be more likely to be directed toward extrapersonal than toward personal space.

Our results clearly show that processing can be less efficient in an unexpected modality than in the same modality when it is expected. Performance depends not only on what actually happens, but also on what is anticipated. This effect of modality expectancy seems primarily to take the form of *costs* for the unexpected modality, rather than benefits for the expected modality (consistent with Pashler's, 1998, account of attentional set, which stresses information exclusion more than information gain). Those *benefits* that are found for the likely target modality seem primarily to be caused by stimulus repetition within that modality, not pure expectancy (compare Tables 1 and 4). Future research could determine exactly where in the processing hierarchy these different effects arise. For instance, ERP measures (see Foxe et al., 1998) and functional imaging (see Driver & Spence, 1998) could be applied to our design, to see whether early sensory responses can be modulated by the modality expectancies and MSEs we have documented here.

## REFERENCES

- BOULTER, L. R. (1977). Attention and reaction times to signals of uncertain modality. *Journal of Experimental Psychology: Human Perception & Performance*, *3*, 379-388.
- CIESIELSKI, K. T., KNIGHT, J. E., PRINCE, R. J., HARRIS, R. J., & HANDMAKER, S. D. (1995). Event-related potentials in cross-modal divided attention in autism. *Neuropsychologia*, *33*, 225-246.
- DAI, H., SCHARF, B., & BUUS, S. (1991). Effective attenuation of signals in noise under focused attention. *Journal of the Acoustical Society of America*, *98*, 798-806.
- DRIVER, J., & SPENCE, C. [J.] (1998). Crossmodal attention. *Current Opinion in Neurobiology*, *8*, 245-253.
- DUNCAN, J. (1980). The demonstration of capacity limitation. *Cognitive Psychology*, *12*, 75-96.
- EIMER, M., & DRIVER, J. (2000). An event-related brain potential study of crossmodal links in spatial attention between vision and touch. *Psychophysiology*, *37*, 697-705.
- EPSTEIN, W., & ROCK, I. (1960). Perceptual set as an artifact of recency. *American Journal of Psychology*, *73*, 214-228.
- FERSTL, R., HANEWINKEL, R., & KRAG, P. (1994). Is the modality-shift effect specific for schizophrenia patients? *Schizophrenia Bulletin*, *2*, 367-373.
- FOXE, J. J., SIMPSON, G. V., & AHLFORS, S. P. (1998). Parieto-occipital ~ 10Hz activity reflects anticipatory state of visual attention mechanisms. *NeuroReport*, *9*, 3929-3933.
- GAWRYSZEWSKI, L. DE G., RIGGIO, L., RIZZOLATTI, G., & UMLTÀ, C. (1987). Movements of attention in the three spatial dimensions and the meaning of 'neutral' cues. *Neuropsychologia*, *25*, 19-29.
- GESCHIEDER, G. A., SAGER, L. C., & RUFFOLO, L. J. (1975). Simultaneous auditory and tactile information processing. *Perception & Psychophysics*, *18*, 209-216.
- KLEIN, R. M. (1977). Attention and visual dominance: A chronometric analysis. *Journal of Experimental Psychology: Human Perception & Performance*, *3*, 365-378.
- KORNBLUM, S. (1967). Choice reaction time for repetitions and non-repetitions: A reexamination of the information hypothesis. *Acta Psychologica*, *27*, 178-187.
- LABERGE, D. (1973). Identification of two components of the time to

- switch attention: A test of a serial and a parallel model of attention. In S. Kornblum (Ed.), *Attention and performance IV* (pp. 71-85). New York: Academic Press.
- MARTIN, M. G. F. (1995). Bodily awareness: A sense of ownership. In J. L. Bermudez, A. Marcel, & N. Eilan (Eds.), *The body and the self* (pp. 267-289). Cambridge, MA: MIT Press.
- PASHLER, H. E. (1998). *The psychology of attention*. Cambridge, MA: MIT Press.
- POSNER, M. I., NISSEN, M. J., & KLEIN, R. M. (1976). Visual dominance: An information-processing account of its origins and significance. *Psychological Review*, **83**, 157-171.
- POSNER, M. I., NISSEN, M. J., & OGDEN, W. C. (1978). Attended and unattended processing modes: The role of set for spatial location. In H. I. Pick & E. Saltzman (Eds.), *Modes of perceiving and processing information* (pp. 137-157). Hillsdale, NJ: Erlbaum.
- POST, L. J., & CHAPMAN, C. E. (1991). The effects of cross-modal manipulations of attention on the detection of vibrotactile stimuli in humans. *Somatosensory & Motor Research*, **8**, 149-157.
- ROLAND, P. E. (1982). Cortical regulation of selective attention in man. A regional cerebral blood flow study. *Journal of Neurophysiology*, **48**, 1059-1078.
- SHIU, L.-P., & PASHLER, H. (1994). Negligible effect of spatial precuing on identification of single digits. *Journal of Experimental Psychology: Human Perception & Performance*, **20**, 1037-1054.
- SPENCE, C. {J.}, & DRIVER, J. (1996). Audiovisual links in endogenous covert spatial attention. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 1005-1030.
- SPENCE, C. {J.}, & DRIVER, J. (1997). On measuring selective attention to an expected sensory modality. *Perception & Psychophysics*, **59**, 389-403.
- SPENCE, C. {J.}, NICHOLLS, M. E. R., GILLESPIE, N., & DRIVER, J. (1998). Cross-modal links in exogenous covert spatial orienting between touch, audition, and vision. *Perception & Psychophysics*, **60**, 544-557.
- SPENCE, C. {J.}, PAVANI, F., & DRIVER, J. (2000). Crossmodal links between vision and touch in covert endogenous spatial attention. *Journal of Experimental Psychology: Human Perception & Performance*, **26**, 1298-1319.
- SPRENGELMEYER, R., LANGE, H., & HÖMBERG, V. (1995). The pattern of attention deficits in Huntington's disease. *Brain*, **118**, 145-152.
- STEIN, J. F. (1989). Representation of egocentric space in the posterior parietal cortex. *Quarterly Journal of Experimental Physiology*, **74**, 583-606.
- STEINMETZ, P. N., ROY, A., FITZGERALD, P. J., HSIAO, S. S., JOHNSON, K. O., & NEIBUR, E. (2000). Attention modulates synchronized neuronal firing in primate somatosensory cortex. *Nature & Neuroscience*, **404**, 187-190.
- ZAHN, T. P., PICKAR, D., & HAIER, R. J. (1994). Effects of clozapine, fluphenazine, and placebo on reaction time measures of attention and sensory dominance in schizophrenia. *Schizophrenia Research*, **13**, 133-144.
- ZUBIN, J. (1975). Problems of attention in schizophrenia. In M. L. Kietzman, S. Sutton, & J. Zubin (Eds.), *Experimental approaches to psychopathology* (pp. 139-166). New York: Academic Press.

## NOTES

1. Note that this analysis of just the crossmodal trials, to reduce the influence of any MSEs associated with ipsimodal trials, assumes that any such MSEs occur only between immediately successive trials. We tested this assumption by examining any effect of whether the target from two trials back was ipsimodal or crossmodal with respect to the current target. An ANOVA (target modality [3] × trial type [ipsimodal vs. crossmodal on two trials back]) on the RT data from the divided-attention blocks (note that there was simply insufficient data to perform a similar analysis on the attend-to-a-modality blocks) revealed no significant effect [ $F(1,19) = 0.9, p = .37$ ] of whether the target two trials back had been ipsimodal ( $M = 495$  msec) or crossmodal ( $M = 491$  msec) and no interaction between trial type and current target modality [ $F(2,38) = 1.2, p = .32$ ]. There was not even any numerical trend in the direction of an MSE for two trials back. Our results thus show that MSEs owing to the modality of a preceding target affect performance only on the *immediately succeeding* trial in the present experiment, which justifies the subanalysis in the main text.

2. We conducted a further analysis of the data to assess whether the 10-msec numerical trends for attentional benefits shown in Table 4 for auditory and visual targets might reach significance in a more powerful analysis that considered only these auditory and visual expected targets relative to the divided-attention baseline: A two-way ANOVA (target modality [auditory vs. visual] × expectancy [expected vs. divided]) again revealed that any benefit from expectancy was far from significant [ $F(1,19) = 1.4, p = .25$ ].

(Manuscript received March 29, 2000;  
accepted for publication June 23, 2000.)