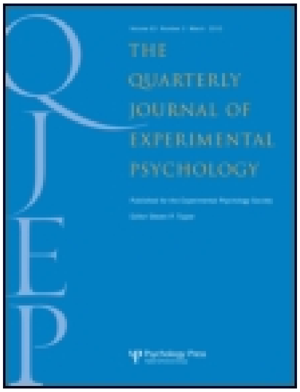

The Quarterly Journal of Experimental Psychology

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pqje20

Modality effects in memory for basic stimulus attributes: A temporal and nontemporal comparison

Ruth S. Ogden [a] [b] & Luke A. Jones [a]

[a] School of Psychological Sciences, University of Manchester, Manchester, UK

[b] School of Natural Sciences and Psychology, Liverpool John Moores University, Liverpool, UK

To cite this article: Ruth S. Ogden & Luke A. Jones (2011): Modality effects in memory for basic stimulus attributes: A temporal and nontemporal comparison, The Quarterly Journal of Experimental Psychology, 64:7, 1354-1371

To link to this article: http://dx.doi.org/10.1080/17470218.2010.543324

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2011, 64 (7), 1354–1371

Psychology Press
Taylor & Francis Group

Modality effects in memory for basic stimulus attributes: A temporal and nontemporal comparison

Ruth S. Ogden[1,2], and Luke A. Jones[1]

[1]School of Psychological Sciences, University of Manchester, Manchester, UK
[2]School of Natural Sciences and Psychology, Liverpool John Moores University, Liverpool, UK

Previous research suggests that there are significant differences in the operation of reference memory for stimuli of different modalities, with visual temporal entries appearing to be more durable than auditory entries (Ogden, Wearden, & Jones, 2008, 2010). Ogden et al. (2008, 2010) demonstrated that when participants were required to store multiple auditory temporal standards over a period of delay there was significant systematic interference to the representation of the standard characterized by shifts in the location of peak responding. No such performance deterioration was observed when multiple visually presented durations were encoded and maintained. The current article explored whether this apparent modality-based difference in reference memory operation is unique to temporal stimuli or whether similar characteristics are also apparent when nontemporal stimuli are encoded and maintained. The modified temporal generalization method developed in Ogden et al. (2008) was employed; however, standards and comparisons varied by pitch (auditory) and physical line length (visual) rather than duration. Pitch and line length generalization results indicated that increasing memory load led to more variable responding and reduced recognition of the standard; however, there was no systematic shift in the location of peak responding. Comparison of the results of this study with those of Ogden et al. (2008, 2010) suggests that although performance deterioration as a consequence of increases in memory load is common to auditory temporal and nontemporal stimuli and visual nontemporal stimuli, systematic interference is unique to auditory temporal processing.

Keywords: Temporal generalization; Nontemporal generalization; Reference memory; Pitch perception; Length perception.

A fundamental question in our understanding of temporal perception is whether there are specialized cognitive processes that enable the perception of time or whether temporal perception is a by-product of other cognitive processes (Wearden, 1999). For example, are there unique properties to the way in which temporal information is stored and maintained in short-term memory (STM) and long-term memory or is temporal information stored in much the same way as other basic stimuli? To answer this question, a small number of previous studies have examined

---

Correspondence should be sent to Ruth S. Ogden, School of Psychology and Natural Sciences, Liverpool John Moores University, Liverpool, L32 AJ, UK. E-mail: r.s.ogden@ljmu.ac.uk

The work reported in this article was conducted as part of the PhD research of the first author whilst at the University of Manchester.

http://www.psypress.com/qjep

DOI:10.1080/17470218.2010.543324

the similarities between temporal processing and other unidimensional processing such as tone pitch and line length (McCormack, Brown, Maylor, Darby, & Green, 1999; McCormack, Brown, Smith, & Brock, 2004; Wearden, Parry, & Stamp, 2002). These studies have generally adopted one of two strategies: either exploring whether the seemingly unique behavioural effects observed in temporal perception can also be replicated using nontemporal stimuli, or examining whether neural activity differs for comparable temporal and nontemporal tasks (Gibbons, Brandler, & Rammsayer, 2003; Maquet et al., 1996; Pouthas, Garnero, Ferrandez, & Renault, 2000).

Experiments examining the behavioural and neural similarities between temporal and nontemporal tasks have typically employed generalization and identification tasks (Gibbons et al., 2003; Maquet et al., 1996; Pouthas et al., 2000). In a typical temporal generalization task, participants are presented with a standard duration, *s*, typically presented as an auditory tone or a visual image on a computer screen or a light-emitting diode (LED). Participants are then presented with a series of comparison durations, which are shorter than, longer than, or equal to the standard. After the presentation of each comparison, participants are asked to indicate whether the comparison was the same duration as the standard or not by pressing Y for yes and N for no. The proportion of yes responses is plotted against the comparison/standard ratio to produce a temporal generalization gradient. Of interest is the location of peak responding, which is the comparison value at which the proportion of yes responses is greatest. In adults this point is usually at the comparison value equal to the standard (Wearden, 1992). Nontemporal generalization tasks are essentially the same as temporal generalization tasks; however, the stimuli vary by some dimension other than duration. Nontemporal stimulus variations include intensity judgements in the form of luminance for lights, or frequency judgements in the form of pitch for tones. In an auditory nontemporal generalization task, for example, participants would receive a standard tone of a given pitch *p* followed by comparison tones of pitches higher than, lower than, or equal to the standard. The task would therefore be to compare the pitch of each comparison with that of the standard pitch.

Temporal versus nontemporal processing: Analysis of neural activity

Comparisons of neural activation during temporal and nontemporal perception indicate that processing differences may exist. Pouthas et al. (2000) used both event-related potential (ERP) and positron emission tomography (PET) analysis to assess performance and activation differences during visual temporal and nontemporal generalization. In the temporal generalization task, comparison durations ranged from 490 ms to 910 ms. In the visual nontemporal task, light luminance was varied. Analysis of the behavioural data showed that performance on the luminance intensity task was superior to performance on the duration task. Network activation in both the duration task and the intensity task was similar; however, there was a difference in the time course of the activation produced by the two tasks, with additional activity being required for the duration task to be completed. In addition to activation in the left frontal and cuneus, temporal processing also produced activation of the right prefrontal area, an area Harrington, Haaland, and Knight (1998) found to be essential to accurate timing.

Distinct neutral activation on comparable temporal and nontemporal tasks was also reported by Gibbons et al. (2003) who recorded ERP activity whilst participants were completing either a temporal or a pitch generalization task. Gibbons employed a shorter duration range in the temporal task than did Pouthas et al. (2000): 125 ms to 275 ms. On both tasks, peak responding occurred at the standard; however, performance was better on the duration task. Analysis of the ERP recordings revealed that in the first 100 ms there was no discernable difference in the ERP recordings from the pitch and temporal generalization tasks. After 100 ms, the recordings from the two tasks were clearly dissociable by the ERP wave forms, suggesting different processing during the two tasks. The wave pattern suggested that there may be an increased role for working memory in the temporal task. Taken together, the results of Gibbons et al. (2003) and Pouthas et al. suggest that on basic psychophysical tasks involving stimulus identification there are some differences in neural activity for temporal and nontemporal processing. It is also noteworthy that these differences occurred despite the differences in the duration range employed in the two studies.

Whilst the findings of Gibbons et al. (2003) and Pouthas et al. (2000) provide illuminating information about the behavioural and neural similarities and differences in temporal and nontemporal processing, they have not looked at precisely how specific processes differ in either form of perception. For example, Gibbons et al. (2003) indicate that temporal processing may require more working memory than nontemporal processing; however, it is unclear how this manifests itself behaviourally. Sceptics may also question whether the knowledge that brain activity differs in temporal and nontemporal tasks advances our understanding of processes involved in temporal perception when the stimuli used are fundamentally different, and methodologies employed (be this the task or the imaging equipment) are often unable to tease out precisely how temporal and nontemporal processing differ and why. To understand the specific ways in which temporal and nontemporal processing differ from one another (rather than just looking at holistic differences in performance), a series of experiments have examined whether seemingly unique properties of temporal perception (typically found in memory operation) are also found in comparable nontemporal tasks.

Temporal versus nontemporal processing: Differences in memory operation

When temporal information is stored in STM and reference memory, seemingly unique effects are frequently reported (Droit-Volet, Clement, & Wearden, 2001; McCormack et al., 1999; McCormack et al., 2004; Ogden, Wearden, & Jones, 2008, 2010; Wearden & Ferrara, 1993; Wearden et al., 2002). When examining STM, Wearden and Ferrara demonstrated that when temporal information is maintained over a short period of delay it decays in a systematic manner whereby the duration stored appears to shrink as the period of delay increases, an effect commonly known as subjective shortening. The existence of subjective shortening in nontemporal tasks was examined by Wearden et al. (2002) by comparing the perception of duration with the perception of the physical length of lines. As in Wearden and Ferrara, participants were presented with a standard, *s*, and a comparison, *c* (a tone of duration *t* in the temporal experiments or a line of length *l* in the nontemporal task), which were separated by a delay of 1 to 10 s. Participants were required to indicate whether the comparison was the same duration as the standard, or shorter or longer than the standard. The results produced using this method are complex, and a full description can be found in Wearden and Ferrara or in Wearden et al. (2002); however, it suffices to say that Wearden et al. (2002) found no evidence of subjective shortening in nontemporal perception. The absence of subjective shortening in nontemporal processing indicates fundamental differences in the way in which temporal and nontemporal information decays in STM.

When examining long-term or reference memory function, Droit-Volet et al. (2001) and McCormack and colleagues (McCormack et al., 1999; McCormack, Brown, Maylor, Richardson, & Darby, 2002; McCormack et al., 2004) have demonstrated that there is a systematic change in the operation of temporal reference memory as we age from childhood to early adulthood and then to old age. Children's performance on temporal generalization tasks is characterized by their tendency to confuse the standard duration more with durations shorter than it than with durations longer than it; this results in leftward skewed gradients. The skewing is thought to occur because children remember the standard as being shorter than it actually was. Timing in the elderly appears to be more variable than timing in younger adults, with temporal generalization gradients appearing to be flatter (see McCormack et al., 1999). McCormack et al. (2002) also suggests that elderly participants may also confuse standard durations with comparisons that are longer than the standard, leading to a right asymmetry.

McCormack et al. (2002) and McCormack et al. (2004) examined whether the systematic age-related changes in reference memory function found in temporal perception would also occur when nontemporal stimuli were stored. McCormack et al. (2004) compared the performance of children and adults on a classic temporal generalization task and on a pitch generalization task. In a further paper, McCormack and colleagues compared elderly and adult temporal and nontemporal perception using an identification task (McCormack et al., 2002). In both studies, comparisons of the performances of the different age groups on the temporal and nontemporal tasks indicated that the systematic age-related changes in reference memory operation were unique to temporal processing and did not extend to similar tasks using nontemporal stimuli.

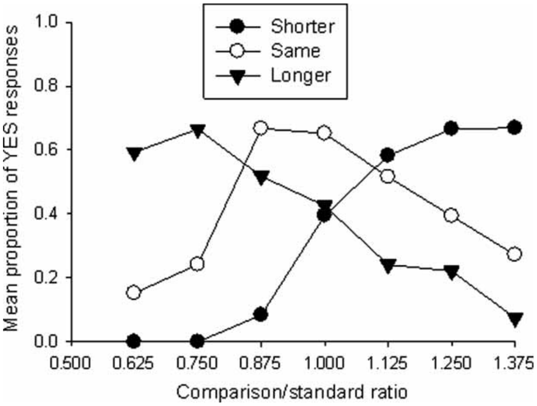

Figure 1. *Temporal generalization gradients from critical comparisons in Experiment 3 of Ogden, Wearden, and Jones (2008). Temporal generalization gradients are shown from the delayed comparisons, from the* shorter, longer, *and* same *groups, demonstrating the shifts in peak responding from the standard. From "The remembrance of times past: Interference in temporal reference memory", by R. S. Ogden, J. H. Wearden, and L. A. Jones, 2008,* Journal of Experimental Psychology: Human Perception and Performance, 34, *p. 1533. Copyright 2008 by the American Psychological Association. Reprinted with permission.*

Modality differences in temporal reference memory: The basis for a new nontemporal comparison

In a recent series of papers, Ogden et al. (2008, 2010) have explored the durability of entries into temporal reference memory. Ogden et al. examined the consequences of encoding, maintaining, and using multiple different durations in reference memory at the same time. A modified temporal generalization method was employed, in which participants learnt a first standard (A) and were then tested on their memory of A (immediate testing of A). Following the immediate testing of A, participants were presented with a second standard (B) and were then tested on their recall of B (immediate testing of B). Following the immediate testing of B, a delay of either 0 (deferred retesting) or 1–30 s (delayed retesting) was interposed; following the delay, participants were retested on their recall of A without A being re-presented. All standard durations ranged from 400 ms to 1,200 ms. Depending on the condition, A was the same duration as B, 400 ms shorter than B, or 400 ms longer than B. Of interest was comparison of performance in immediate testing of A with that of deferred and delayed retesting of A. The results were surprising; when all standards and comparisons were auditory stimuli (500-Hz tones), extreme systematic interference effects were reported in some conditions (see Figure 1). When A was less than B, peak responding shifted from the standard in immediate testing to the shortest of the comparison durations in delayed retesting, meaning that in Ogden et al's. (2008) Experiment 2, for example, 82% of participants believed that the standard was 37.50% shorter than it actually was. When A was greater than B, peak responding shifted from the standard in immediate testing to the longest of the comparison durations in delayed retesting.[1] When A and B

---

[1] Although temporal generalization functions in humans always exhibit right asymmetry, the systematic interference effect when present in the rightward direction massively exceeds this natural asymmetry, and importantly the effect can be obtained in both directions—that is, the effect is manifest in the opposite direction to the natural asymmetry when the memory of the standard is systematically shortened.

were the same duration, no shifts in the location of peak responding were observed. The results indicate that reference memory is unable to accurately maintain two different auditory temporal standards at the same time. When required to do so, the memory of the first standard becomes systematically distorted by the presence of the second standard, leading to shifts in the location of peak responding.

The extreme nature of the systematic distortion observed in Ogden et al. (2008) led the authors to explore whether similar effects would be found if A and B were both visual stimuli (blue squares; Ogden et al., 2010). Comparing inference effects when encoding different modalities would clarify whether temporal reference memory operates in the same way for stimuli of different modalities. As in Ogden et al. (2008), when A equals B no evidence of interference was reported. Unlike in Ogden et al. (2008), however, no interference was reported when A was less than B or A was greater than B (400-ms difference). The absence of any interference to the memory of A after encoding and maintaining B indicated that visual entries in temporal reference memory may be more durable than auditory entries into reference memory. The modality differences reported in Ogden et al. (2008, 2010) indicate that there may be fundamental differences in the way in which auditory and visual temporal information is stored and maintained in reference memory, with potentially unique systematic interference effects occurring when multiple different auditory durations are stored.

Modality effects in memory are not unique to temporal perception having been consistently observed in verbal short-term memory recall tests for the last 40 years. Experiments typically report that recall of verbal information (in particular the recency effect) is superior if it is presented auditorily (spoken) rather than visually (written; Murdock & Walker, 1969). However this superiority can be reduced or eliminated by postlearning distraction (Nairne & Walters, 1983). Modality effects also appear to exist in long-term memory, albeit in a more complex manner (Conway & Gathercole, 1987). For example, Conway and Gathercole reported that under some conditions memory for words was superior when auditory presentation was used. However, Engle and Mobley (1976) reported situations in which words presented visually were better recalled than those using auditory presentation. Similarly, some studies report no modality differences in long-term memory recognition when longer retention delays are employed (e.g., Kirsner, 1974; Lehman, 1982). Exploring modality differences in memory operations using verbal stimuli appears to produce complex patterns of results. The present study therefore aimed to explore modality differences in reference memory operation using nonverbal stimuli.

The current study explored whether the systematic interference observed when multiple auditory temporal standards are encoded into reference memory is a unique feature of the way in which temporal information is stored in reference memory or whether it is a more general characteristic of the storage of auditory stimuli per se. The study also explores whether visual superiority is also a feature of nontemporal entries into reference memory. The method developed in Ogden et al. (2008) was employed using nontemporal stimuli. As an auditory comparison to Ogden et al. (2008), participants were required to encode and maintain multiple pitches, and as a visual comparison to Ogden et al. (2010), participants were required to encode and maintain multiple lines of different physical lengths. The interference effects obtained using auditory and visual nontemporal stimuli were compared to those reported using auditory and visual temporal stimuli in Ogden et al. (2008, 2010).

EXPERIMENT 1

Ogden et al. (2008, 2010) have demonstrated that auditory temporal reference memory has a limited capacity and as a result can only store one item veridically across a period of delay. Ogden et al. reported that when multiple auditory temporal standards of different durations are stored in reference memory over a period of delay, identification of the first encoded standard was poor, with peak responding shifting systematically from the standard in immediate testing to the shortest or longest of the comparison durations in delayed retesting (see Ogden et al., 2008, Experiment 3). Experiment 1 therefore examined whether the systematic interference effects observed in Ogden et al. (2008, 2010) would also occur when participants were required to encode multiple auditory nontemporal standards. The modified temporal generalization task developed in Ogden et al. (2008) was employed; however, the pitch of the standards and comparisons varied instead of their duration. Two experiments (an easy comparison/standard spacing ±10 Hz and a hard comparison/standard spacing ±20 Hz) were conducted to explore potential changes in performance at different levels of task difficulty. If encoding and maintaining multiple different pitches over a period of delay have no effect on generalization performance (i.e., gradients at delayed retest do not appear flatter than those in immediate testing), then this would indicate that auditory temporal entries into reference memory are stored in a different way from auditory entries in general. It would also indicate that the way in which auditory temporal information is encoded and maintained is more vulnerable to interference than are other auditory entries. If, however, encoding and maintaining multiple pitches over a period of delay result in significant interference (systematic or otherwise) in generalization performance at retest, this would indicate that auditory temporal entries into reference memory are no more vulnerable to interference than are other basic auditory entries into reference memory.

Method

Participants

Fifty-six University of Manchester & Liverpool John Moores University students participated for course credit. Participants were allocated to one of four testing groups (*easy B higher*, 12 participants; *easy same*, 12 participants; *hard B higher*, 16 participants; *hard same*, 16 participants).

Stimuli and apparatus

The experimental stimuli were presented on a Dell Optiplex GX620 computer. E-prime software was used to write the program and to control and record all experimental events.

Procedure

Participants received all trials in one session. A modified temporal generalization method was used in which participants had to compare the pitch of different standards to that of different comparisons. All standards and comparisons were 200-ms tones. The following frequencies were each used once as the first standard (A): 330 Hz, 340 Hz, 350 Hz, 360 Hz, 370 Hz, 380 Hz, 390 Hz, 400 Hz, 410 Hz, 420 Hz, 430 Hz, 440 Hz, 450 Hz, 460 Hz, 470 Hz; the order of presentation was randomized for each participant. The second standard (B) was always 300 Hz higher in pitch than A.

In all trials the first standard (A) was presented three times followed by eight comparison tones. A delay imposed between each presentation of the standard was drawn at random from a uniform distribution ranging from 1,500–2,000 ms. In the easy condition, comparison tones were the frequency of A minus 20 Hz, 40 Hz, and 60 Hz; A (presented twice to avoid participants continually responding “no”); and A plus 20 Hz, 40 Hz, and 60 Hz. In the hard condition, comparison tones were the frequency of A minus 10 Hz, 20 Hz, and 30 Hz; A (presented twice to avoid participants continually responding “no”); and A plus 10 Hz, 20 Hz, and 30 Hz (the same comparison spacing was used as that in Gibbons et al., 2003). Comparisons were presented in a random order. After the presentation of each comparison, participants were asked whether it was the same pitch as A (*immediate testing of A*). Participants were then prompted to press Y for yes and N for no. After each response, participants pressed the spacebar to receive the next comparison.

Following the response to the final comparison for A, the second standard, B, was presented three times. In the *same* group, A and B were the same pitch. In the *B higher* group, the standard B was 300 Hz higher in frequency than A. The presentation of B was also followed by eight comparisons. In the easy condition, comparison tones were the frequency of B minus 20 Hz, 40 Hz, and 60 Hz; B (presented twice to avoid participants continually responding "no"); and B plus 20 Hz, 40 Hz, and 60 Hz. In the hard condition, comparison tones were the frequency of B minus 10 Hz, 20 Hz, and 30 Hz; B (presented twice to avoid participants continually responding "no"); and B plus 10 Hz, 20 Hz, and 30 Hz. After each comparison, participants were asked whether the comparison was the same pitch as B *(immediate testing of B)*. Participants were prompted to press Y if the pitch was the same as B or N if it was different. After each response, participants pressed the spacebar to receive the next comparison.

Following the response to the final comparison for B, a delay of 0 *(deferred retesting)*, 5, 10, 20, or 30 *(delayed retesting*) seconds was imposed. Following this delay, participants were presented with the same eight comparisons as those used in the immediate testing of A. Comparisons were presented in a random order. After the presentation of each comparison, participants were asked whether it was the same pitch as the first standard. Participants were prompted to press Y for yes and N for no. After each response, participants pressed the spacebar to receive the next comparison. No performance feedback was given. Each delay duration was used three times, giving a total of 15 trials. See Figure 2 for the method in schematic form.

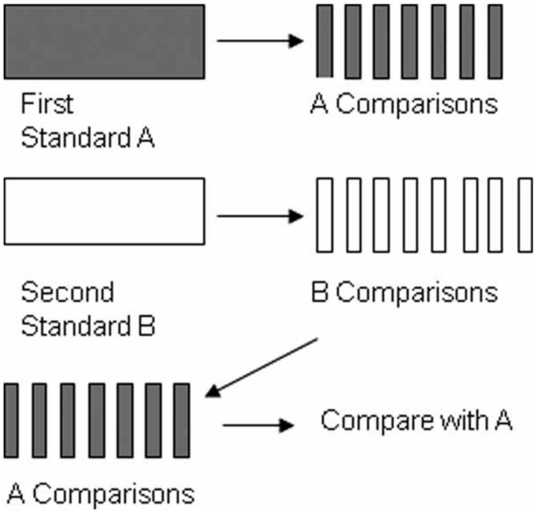

Figure 2. *Method in schematic form.*

Results

Easy condition

Figure 3 shows pitch generalization gradients of the mean proportion of yes responses plotted against the comparison/standard ratio. The upper panel shows data from the *B higher* testing group; the lower panel shows data from the *same* group. Data from all testing conditions (immediate, deferred, and delayed) are shown. Inspection of Figure 3 shows that in immediate testing performance is good, and peak responding occurs at the standard. In the deferred and delayed testing conditions, gradients are flatter with fewer correct identifications of the standard. In some cases there also appears to be a shift in the location of peak responding.

An overall repeated measures analysis of variance (ANOVA) used testing condition (immediate, deferred, and delayed testing) and comparison/standard spacing as within-subject factors and group (*same* or *B higher*) as a between-subjects factor. There was a significant effect of comparison/standard spacing, $F(6, 144) = 39.59$, $p < .001$, $\eta^2 = .62$, and a significant interaction between comparison/standard spacing and testing condition, $F(30, 720) = 2.26$, $p < .001$, $\eta^2 = .09$. There was no significant effect of testing condition, $F(5, 120) = 1.21$, $p = .31$, or group, $F(1, 24) = 1.38$, $p = .25$. There were no other significant two- or three-way interactions.

Individual ANOVAs were then conducted for each testing group (*same* and *B higher*), with testing condition (immediate, deferred, and delayed testing) and comparison/standard spacing as within-subject factors. For the *B higher* group there was a significant effect of comparison/standard ratio, $F(6, 78) = 18.38$, $p < .001$, $\eta^2 = .59$. There was no significant effect of length of delay, $F(5, 65) = 1.05$, $p = .39$; however, there was a significant interaction between length of delay and comparison/standard ratio, $F(30, 390) = 1.50$,

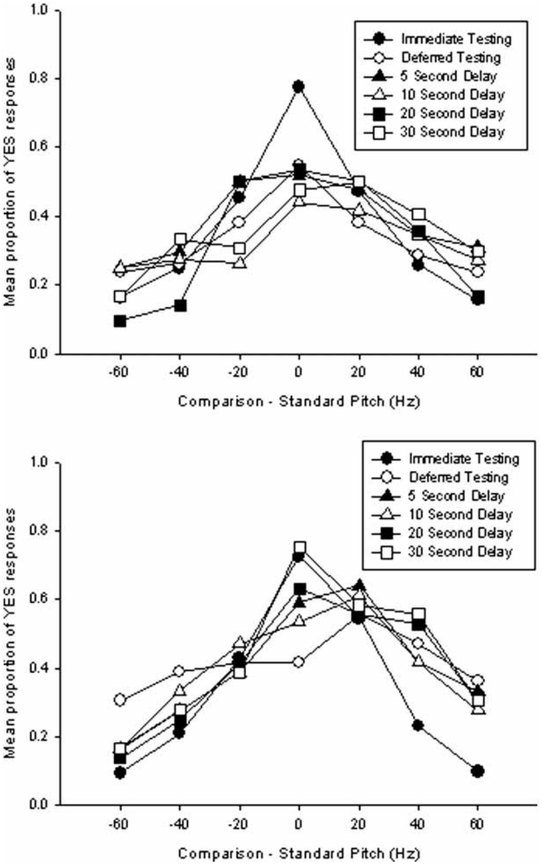

Figure 3. *Pitch generalization gradients—proportion of yes responses (identifications of a comparison pitch as the standard) plotted against comparison/standard ratio—from the easy condition of Experiment 1. Upper panel:* B higher. *Lower panel:* same. *Within each panel, data from the immediate, deferred, and delayed comparisons are presented.*

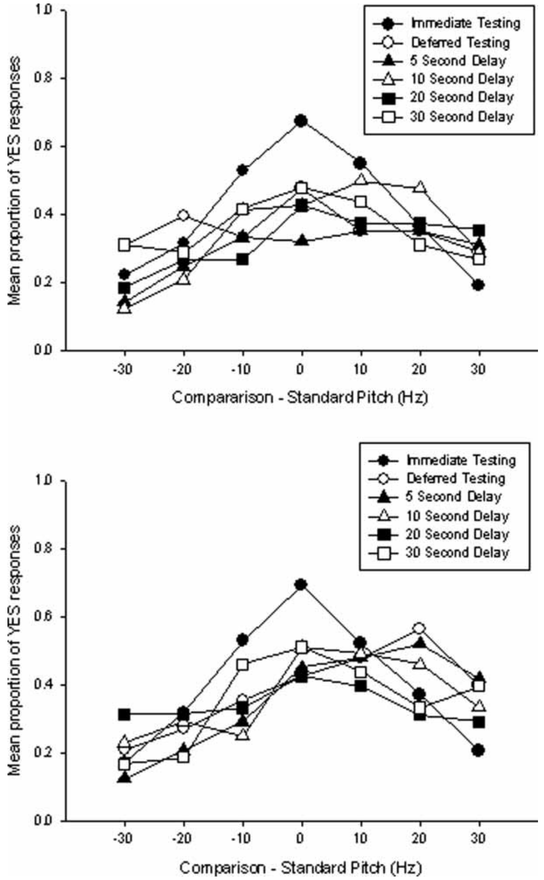

Figure 4. *Pitch generalization gradients—proportion of yes responses (identifications of a comparison pitch as the standard) plotted against comparison/standard ratio—from the hard condition of Experiment 1. Upper panel:* B higher. *Lower panel:* same. *Within each panel, data from the immediate, deferred, and delayed comparisons are presented.*

$p < .05$, $\eta^2 = .10$. For the *same* condition there was a significant effect of comparison/standard ratio, $F(6, 66) = 21.47$, $p < .001$, $\eta^2 = .67$. There was no significant effect of length of delay, $F(5, 55) = 2.04$, $p = .09$; however, there was a significant interaction between length of delay and comparison/standard ratio, $F(30, 330) = 1.73$, $p < .05$, $\eta^2 = .14$.

Hard condition

Figure 4 shows pitch generalization gradients of the mean proportion of yes responses plotted against the comparison/standard ratio. The upper panel shows data from the *B higher* testing group; the lower panel shows data from the *same* group. Data from all testing conditions (immediate, deferred, and delayed) are shown. Inspection of Figure 4 shows that in immediate testing, performance is good, and peak responding occurs at the standard. In the deferred and delayed testing conditions, gradients are flatter, suggesting relatively little sensitivity to the length of comparisons or a poor memory of the duration of the standard. The flatness of the gradients makes it difficult to ascertain a clear peak in responding in the deferred and delayed conditions; however, in many cases it appears to be at durations other than the standard. All suggestions were supported by statistical analysis.

An overall repeated measures ANOVA used testing condition and comparison/standard spacing as within-subject factors and group as a between-subjects factor. There was a significant effect of comparison/standard spacing, $F(6, 180) = 19.46$, $p < .01$, $\eta^2 = .39$, and a significant effect of testing condition, $F(5, 150) = 4.00$, $p < .01$, $\eta^2 = .12$. There was also a significant interaction between comparison/standard spacing and testing condition, $F(30, 900) = 1.06$, $p < .01$, $\eta^2 = .08$. There was no significant effect of group, $F(1, 30) = 0.21$, $p = .65$, nor were there any other significant two- or three-way interactions.

Individual ANOVAs were then conducted for each testing group with testing condition and comparison/standard spacing as within-subject factors. For the *B higher* group, there was a significant effect of comparison/standard spacing, $F(6, 90) = 8.23$, $p < .001$, $\eta^2 = .35$, and a significant effect of testing condition, $F(5, 75) = 2.59$, $p < .05$, $\eta^2 = .15$. There was also a significant interaction between comparison/standard spacing and testing condition, $F(30, 450) = 1.67$, $p < .05$, $\eta^2 = .10$. To confirm no additional effect of increased delay on performance, a further ANOVA was conducted using only data from the deferred and delayed conditions. There was a significant effect of comparison/standard spacing, $F(6, 90) = 5.16$, $p < .01$. There was no significant effect of testing condition, $F(4, 60) = 1.35$, $p = .26$, nor was there a significant interaction between comparison/standard spacing and testing condition, $F(24, 360) = 1.02$, $p = .45$, indicating that there was no significant difference between performance on the deferred and delayed testing conditions.

For the *same* group there was a significant effect of comparison/standard spacing, $F(6, 90) = 11.77$, $p < .01$, $\eta^2 = .44$, and a significant interaction between comparison/standard spacing and testing condition, $F(30, 450) = 1.87$, $p < .01$, $\eta^2 = .11$. There was no significant effect of testing condition, $F(5, 75) = 1.86$, $p = .11$. To confirm no additional effect of increased delay on performance, a further ANOVA was conducted using only data from the deferred and delayed conditions. There was a significant effect of comparison/standard spacing, $F(6, 90) = 8.42$, $p < .001$. There was no significant effect of testing condition, $F(4, 60) = 0.90$, $p = .47$, nor was there a significant interaction between comparison/standard spacing and testing condition, $F(24, 360) = 1.01$, $p = .34$, confirming that there was no significant difference between performance on the deferred and delayed testing conditions.

To confirm that the absence of a systematic shift in the location of peak responding in Experiment 1 was not due to differing levels of task difficulty between the temporal (Ogden et al., 2008, 2010) and nontemporal tasks, we compared the proportion of hits (correct identifications of the standard) and correct rejections (no responses to comparisons that were not the same as the standard) during the immediate testing of A on the temporal and nontemporal tasks.[2] Temporal data were taken from Ogden et al. (2008, Experiment 4a); this experiment was selected because it employs the same delay durations as those in the current experiment. Comparison of responding during immediate testing on the easy pitch generalization task and temporal generalization task indicated significant differences in task difficulty. Independent-samples *t* tests showed no significant difference in the proportion of hits, $t(31) = -0.20$, $p = .84$ (pitch $M = .78$, temporal $M = .76$) across the two tasks. There were, however, significant differences between the proportion of correct rejections $t(31) = -2.52$, $p < .05$ (pitch $M = .71$, temporal $M = .59$), indicating that the easy pitch task was significantly easier than the temporal task. Comparison of responding during the hard pitch generalization task and the temporal generalization task indicated comparable levels of task difficulty. Independent-samples *t* tests showed no significant difference between the proportion of

---

[2] Note that only data from the *B higher* condition were analysed, as Ogden et al.'s (2008) Experiment 4 did not employ a *same* condition.

hits, $t(33) = 1.55, p = .13$ (pitch $M = .68$, temporal $M = .76$) or correct rejections, $t(33) = -1.11, p = .27$ (pitch $M = .64$, temporal $M = .59$) for the hard pitch generalization experiment and Experiment 4a of Ogden et al. (2008). This suggests that the absence of a systematic shift in the location of peak responding in the pitch experiments did not occur because the pitch task was easier than the duration task.

Discussion

Examination of Figures 3 and 4 shows that in immediate testing performance was good, and peak responding occurred at the standard, indicating that the standard had been stored veridically. Regardless of the level of task difficulty or the relationship between the first and second standards (e.g., *same* pitch or *B higher*), correct identifications of the standard were significantly reduced after the second standard had been encoded (deferred and delayed retesting). Performance deterioration appears to be worse in the hard condition where responding appears to be almost at random in some testing conditions, with no discernable peak. These flattened gradients suggest that encoding and maintaining a second standard in reference memory result in significant interference to the memory of the first standard, even when the first and second standards were of the same pitch. The inference resulting from encoding the second standard is so severe that in the hard condition the memory of the first standard becomes too degraded to be useful.

Comparison of the results of Experiment 1 with the temporal data from Ogden et al. (2008, 2010) suggests that severe deterioration in generalization performance after encoding multiple different auditory standards is not unique to temporal stimuli. In both temporal and nontemporal generalization, when multiple different standards are encoded and maintained over a period of delay, the memory of the first standard suffers interference as a result of encoding the second standard. In temporal generalization, this inferences causes a systemic shift in the location of peak responding, presumably due to participants constructing a new, albeit inaccurate, representation of the standard; however, in pitch generalization, responding simply becomes more variable in a nonsystematic manner. Comparison of the temporal and nontemporal data suggests that inference appears to be greater in pitch generalization than in temporal generalization. For example, in Ogden et al. (2008), performance deterioration at retest was only observed (a) when A and B were of different durations and (b) when A and B had to be maintained over a period of delay of at least one second. In the pitch generalization experiments, however, significant performance deterioration was observed when A and B were the same pitch and during deferred retesting rather than just delayed retesting. Despite these differences, the deterioration observed on both temporal and nontemporal tasks suggests that the capacity limitations for auditory temporal stimuli observed by Ogden et al. (2008, 2010) are not unique to temporal entries into reference memory but appear also to be a feature of representations of pitch information. Crucially, however, it is only on temporal tasks that this deterioration is systematic; on nontemporal tasks, responding simply becomes generally more variable in a nonsystematic manner, regardless of the level of task difficulty. Therefore, unlike in auditory temporal generalization where participants appear to construct a new standard for use in delayed retesting by using a rule such as "I can't remember A but I remember that A was much shorter/longer than B so I will respond yes at durations that are much shorter/longer than B", in auditory nontemporal generalization participants just use their decayed memory of A to complete the task, rather than their memory of the relationship between A and B. In the temporal task, this results in systematic distortion to the memory of the first standard; however, in the nontemporal task, the responding suggests little or no recognition of the first standard rather than a systematic distortion of it.

Whilst reference memory (temporal or nontemporal) appears unable to veridically store more than one auditory item (temporal or nontemporal) at a time, Ogden et al. (2010) demonstrated that the same is not necessarily true for visual stimuli. Ogden et al. suggest that fundamental differences appear to exist between the capacity of reference memory for auditory and visual temporal stimuli; whilst encoding multiple auditory temporal standards leads to significant systematic performance deterioration, encoding multiple visually presented durations does not. This suggests that visual representations of duration in reference memory are more resistant to interference than are auditory representations. Given the known modality differences in reference memory function when storing temporal information, Experiment 2 explored whether increased durability of visual representations in reference memory is a unique feature of the way in which temporal information is stored or whether nontemporal visual entries are also less vulnerable to interference or overwriting than are auditory entries.

EXPERIMENT 2

Experiment 2 examined the effect of encoding multiple visual nontemporal standards on generalization performance. According to Ogden et al. (2010), encoding and storing multiple visual temporal standards of different durations do not lead to systematic interference in performance (or even significant performance detriment at delayed retesting). In Experiment 2, therefore, an absence of systematic interference as a result of encoding a second visual standard would indicate similarities between reference memory operation for visual temporal and nontemporal information. An absence of interference would also confirm Ogden et al.'s (2010) suggestion that there are differences in the operation of reference memory for stimuli of different modalities with visual representations appearing to be less vulnerable to interference and overwriting than auditory representations. The presence of systematic interference or significantly more variable performance after encoding a second standard would indicate differences in the way that visual temporal and nontemporal information is encoded and maintained in reference memory.

To examine the effect of encoding multiple visual nontemporal standards into reference memory, the modified generalization method used in Experiment 1 was employed; however, in this task participants were required to compare the physical length of standard lines to that of comparison lines. Line length has previously been employed as a comparison to temporal perception in Elvevag, Brown, McCormack, Vousden, and Goldberg (2004) and Wearden et al. (2002). As in Experiment 1, two levels of difficulty were employed.

Method

Participants

Fifty-six University of Manchester and Liverpool John Moores University students participated for course credit. Students were allocated to one of four testing groups (*easy same*, 15 participants; *easy A longer*, 15 participants; *hard same*, 13 participants; and *hard A longer*, 13 participants).

Procedure

The procedure was essentially the same as that in Experiment 1; however, physical line length was used as the comparison test rather than pitch. All standards and comparisons were presented as vertical yellow lines on a black background. All first standard (A) lengths were drawn from a uniform distribution ranging from 5 cm to 10 cm in length. Between the presentations of each standard there was a delay drawn randomly from a uniform distribution ranging from 1,500 ms to 2,000 ms. In the *same* testing group, A was the same length as B; in the *A longer* testing group, A was 30% longer than B. In the *easy* condition, the comparisons for A and B were the standard multiplied by 0.25, 0.50, 0.75, 1.00, 1.25, 1.50, 1.75 (the same comparison/standard ratios as those in Ogden et al., 2010). In the hard condition, the comparisons for A and B were the standard multiplied by 0.625, 0.750, 0.875, 1.00, 1.125, 1.250, 1.375. Durations of delay prior to retesting remained the same as those in Experiment 1 (0, 5, 10, 20, and 30 s), and each one was used three times. No performance feedback was given. All apparatus remained the same as that in Experiment 1.

Results

Easy condition

Figure 5 shows line generalization gradients: the mean proportion of yes responses plotted against comparison/standard ratio for the two testing conditions, *A longer* (upper panel) and *same* (lower panel). Within each panel are data from immediate, deferred, and delayed testing conditions. Examination of Figure 5 shows that in the *same* condition peak responding occurred at the standard in all conditions, and there is no consistent flattening of the gradients as the period of delay prior to retesting increases. In the *A longer* condition, peak responding is at the standard in all conditions except deferred. Gradients appear to get progressively flatter as the delay to retesting increases.

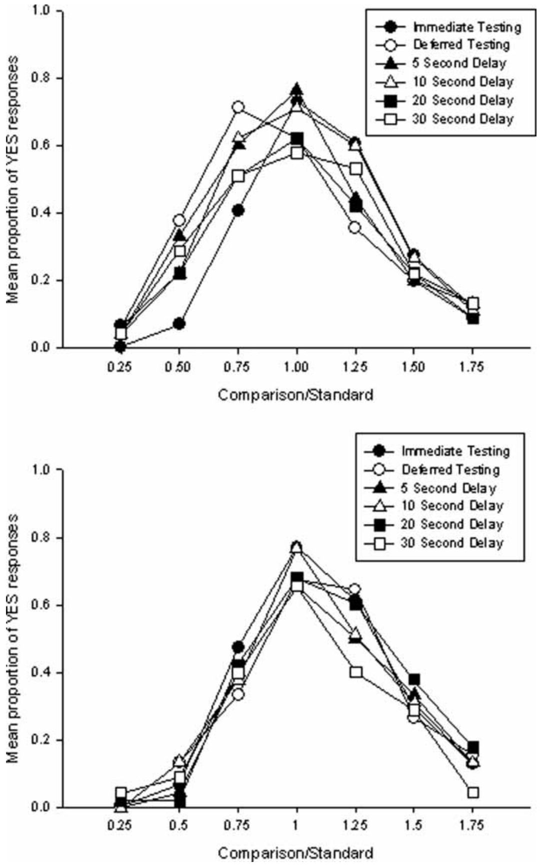

Figure 5. *Line generalization gradients—proportion of yes responses (identifications of a comparison line as the standard) plotted against comparison/standard ratio—from the easy condition of Experiment 2. Within each panel, data from the immediate, deferred, and delayed comparisons are presented. Upper panel:* A longer *testing group. Lower panel:* same *testing group.*

An overall ANOVA conducted on all data from both testing groups (*same* or *A longer*) found a significant effect of comparison/standard ratio, $F(6, 168) = 81.09, p < .001, \eta^2 = .74$, and significant two-way interactions between comparison/standard ratio and testing group, $F(6, 168) = 3.55$, $p < .01$, $\eta^2 = .11$, and length of delay and testing group, $F(5, 140) = 2.50$, $p < .05$, $\eta^2 = .09$. There was also a significant three-way interaction between comparison/standard ratio, length of delay, and testing condition, $F(30, 840) = 1.71, p < .05, \eta^2 = .06$. There was no significant effect of length of delay alone, $F(5, 140) = 1.33$, $p = .26$, nor of testing group, $F(1, 28) = 1.66$, $p = .21$. There was also no interaction between comparison/standard ratio and length of delay, $F(30, 840) = 1.21, p = .20$.

Individual ANOVAs for each testing group revealed that for the *same* condition there was a significant effect of comparison/standard ratio, $F(6, 84) = 49.34$, $p < .001$, $\eta^2 = .78$. There was no significant effect of length of delay, $F(5, 70) = 2.08$, $p = .11$, nor was there a significant interaction between length of delay and comparison/standard ratio, $F(30, 420) = 0.99$, $p = .47$. For the *A longer* condition, there was a significant effect of comparison/standard ratio, $F(6, 84) = 36.38$, $p < .001$, $\eta^2 = .72$. There was no significant effect of length of delay, $F(5, 70) = 1.78$, $p = .13$; however, there was a significant interaction between length of delay and comparison/standard ratio, $F(30, 420) = 1.82$, $p < .01$, $\eta^2 = .12$.

Hard condition

Figure 6 shows line generalization gradients: the mean proportion of yes responses plotted against

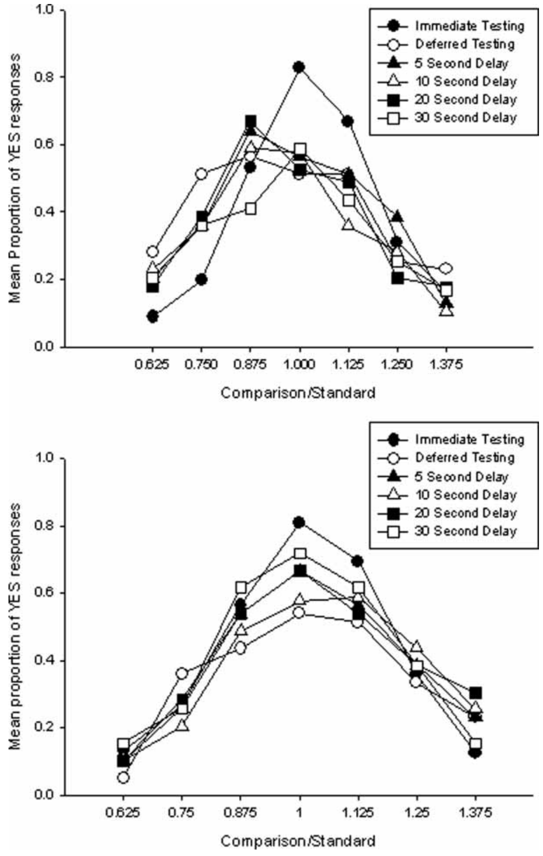

Figure 6. *Line generalization gradients—proportion of yes responses (identifications of a comparison line as the standard) plotted against comparison/standard ratio—from the hard condition of Experiment 2. Within each panel, data from the immediate, deferred, and delayed comparisons are presented. Upper panel:* A longer *testing group. Lower panel:* same *testing group.*

comparison/standard ratio for the two testing conditions, *A longer* (upper panel) and *same* (lower panel). Within each panel are data from immediate, deferred, and delayed testing conditions. Examination of Figure 6 shows that in both conditions correct identification of the standard is good; however, in deferred and delayed retesting, gradients become flatter, and responding appears more variable. In the *same* condition, peak responding remains at the standard in the deferred and delayed testing conditions; however, in the *A longer* condition, peak shifts to comparison durations shorter than the standard (note that this shift is in the opposite direction to that predicted by Ogden et al., 2010).

An overall ANOVA conducted on all data found a significant effect of comparison/standard ratio, $F(6, 138) = 37.07, p < .001, \eta^2 = .62$, and significant two-way interactions between comparison/standard ratio and testing group (*same* or *A longer)*, $F(6, 138) = 2.61, p < .01, \eta^2 = .10$, and length of delay and comparison/standard ratio, $F(30, 690) = 2.81, p < .001, \eta^2 = .09$. There was no significant effect of length of delay ($F < 1$) or testing condition ($F < 1$). There were also no other significant two- or three-way interactions.

Individual ANOVAs conducted for each testing group revealed that for the *same* condition there was a significant effect of comparison/standard ratio, $F(6, 72) = 21.08, p < .001, \eta^2 = .64$. There was no significant effect of length of delay, $F(5, 60) = 1.43, p = .25$, nor was there a significant interaction between length of delay and comparison/standard ratio, $F(30, 360) = 0.93, p = .57$. For the *A longer* condition there was a significant effect of comparison/standard ratio, $F(6, 66) = 19.55, p < .001, \eta^2 = .64$. There was no significant effect of length of delay, $F(5, 55) = 1.05, p = .42$; however, there was a significant interaction between length of delay and comparison/standard ratio, $F(30, 330) = 2.40, p < .001, \eta^2 = .18$.

As in Experiment 1, to ensure comparable levels of task difficulty between the experiments reported here and the temporal tasks reported in Ogden et al. (2010, Experiment 2) we compared the proportion of hits and correct rejections across the tasks. Comparison of performance during immediate testing in the easy condition of our Experiment 2 and of Ogden et al.'s (2010) Experiment 2 indicated that the easy line length task was easier than the temporal task. Independent-samples *t* tests showed no significant difference in the proportion of hits, $t(55) = 0.63, p = .53$ (line length $M = .78$, temporal $M = .75$); however, there was a significant difference in the proportion of correct rejections, $t(55) = -5.56, p < .001$ (line length $M = .74$, temporal $M = .61$). Comparison of the performance during immediate testing in the hard condition of Experiment 2 (nontemporal) and of Ogden et al.'s (2010) Experiment 2 (temporal) indicates comparable levels of task difficulty. Independent-samples t tests showed no significant difference in the proportion of hits, $t(51) = 1.09$, $p = .28$ (line length $M = .82$, temporal $M = .78$), or correct rejections $t(51) = -1.78$, $p = .08$ (line length $M = .66$, temporal $M = .61$). This suggests that when the difficulty of the temporal and nontemporal tasks is comparable, encoding multiple standards has a more detrimental effect on generalization performance in a nontemporal task than in a temporal task.

Cross-modal comparisons

To ensure that performance differences on the pitch and line length tasks were not due to differences in their difficulty, we compared the proportion of hits and correct rejections during immediate testing on both tasks. Comparison of responding on the easy pitch and the easy line length experiment indicated comparable levels of difficulty on the two tasks. Independent-samples t tests showed no significant differences in the proportion of hits, $t(54) = -0.05$, $p = .96$ (line length $M = .75$, pitch $M = .75$), or correct rejections, $t(54) = 0.36$, $p = .93$ (line length $M = .74$, pitch $M = .72$). Comparison of responding during the hard pitch and line length tasks indicated that there were significant differences in the difficulty of the tasks. Independent-samples t tests showed significant differences in the proportion of hits, $t(53) = -3.43$, $p < .001$, with significantly more hits in the line length task ($M = .82$) than in the pitch task ($M = .68$). There was no significant difference in the proportion of correct rejections, $t(53) = -0.59$, $p = .55$ (line length $M = .66$, pitch $M = .64$).

Finally, we also examined whether deteriorations in correct identifications of the standard were comparable between the immediate and delayed (5, 10, 20, and 30 s) conditions in the visual and auditory conditions. For the easy experiments (pitch and line length), independent-samples t tests showed a significant difference in deteriorations in correct identifications of the standard for the different conditions, $t(27) = 3.69$, $p < .001$, with a greater deterioration in the pitch experiment ($M = .28$) than in the line length experiment ($M = .06$). For the *same* condition, independent-samples t tests showed no significant difference in deteriorations in correct identifications of the standard for the different conditions, $t(25) = 1.09$, $p = .29$ (pitch $M = .14$, line length $M = .08$). For the hard conditions, independent-samples t tests showed no significant difference in deteriorations in correct identifications of the standard for the *same* conditions (pitch and line length), $t(27) = 1.05$, $p = .30$ (pitch $M = .22$, line length $M = .15$) or the *different* conditions, $t(27) = -0.56$, $p = .96$ (pitch $M = .26$, line length $M = .26$). This indicates that under hard task conditions, encoding multiple standards and maintaining them over a period of delay have a comparable effect on delayed correct identifications of the standard in both modalities.

Discussion

The results of Experiment 2 demonstrate that when A and B are both visual and are the same length there is little or no deterioration in generalization performance after encoding the second standard. When A and B are different lengths, there appears to be some deterioration in generalization performance after encoding the second standard, as gradients are flatter in the delayed retesting conditions; this is true to a greater extent in the hard condition. In the easy task for both the *same* and the *A longer* conditions, there was no evidence of systematic interference characterized by shifts in the location of peak responding as observed using auditory temporal stimuli in Ogden et al. (2008). In the hard task, the location of peak responding did appear to shift leftward; however, this is not consistent over all conditions of delayed retesting, and the direction of the shift is the opposite to that observed in Ogden et al. (2008). The absence of systematic interference characterized by a shift in the location of peak responding may seem somewhat unsurprising considering the absence of any systematic interference in the visual temporal version of this task

(Ogden et al., 2010, Experiment 2). The deterioration in performance over increasing delays and leftward shifts in peak responding in some conditions in the *A longer* condition does, however, suggest that the second standard did influence the memory of the first standard by making it more variable.

Comparison of the deterioration in correct identifications of the standard between immediate testing and delayed retesting in the pitch and line length experiments indicates that under easy task conditions there is significantly less interference when the two standards are visual than when they are auditory. This amplified vulnerability of auditory nontemporal information to inference in reference memory mirrors the finding that auditory temporal information is also more vulnerable to interference than is visual temporal information; however, it should be noted that once task difficulty increases, performance deterioration is comparable across the two modalities (hard conditions) for nontemporal stimuli.

GENERAL DISCUSSION

The two experiments reported in this article aimed to further explore whether temporal processing is dependent on specialized processes or whether it uses the same cognitive processes as other psychophysical judgements. As in previous research (McCormack et al., 1999; McCormack et al., 2004; Wearden & Ferrara, 1993; Wearden et al., 2002), we explored whether seemingly unique effects observed in temporal perception research also occur when stimuli vary by some dimension other than duration. Ogden et al. (2008, 2010) reported that our ability to store multiple temporal standards over a period of delay differed depending on the modality in which the stimuli were presented. Whilst participants were able to accurately maintain multiple visual temporal standards over a period of delay, they appeared unable to accurately maintain multiple different auditory temporal standards over a period of delay. When they were required to maintain multiple auditory standards of different durations over a period of delay, their ability to accurately identify the standard was significantly reduced. Analysis demonstrated that as a consequence of encoding multiple different standards over a period of delay, there was a systematic distortion to the participant's reference memory representation of the standard, leading to shift in the location of peak responding. The direction of the shift was dependent on the relationships between the durations encoded. Taken together, the results of Ogden et al. (2008, 2010) suggest that reference memory has a greater capacity for temporal information when it is presented in the visual modality than when it is presented in the auditory modality. The current study therefore sought to examine whether nontemporal auditory stimuli (tones varying by pitch) were also more vulnerable to inference than nontemporal visual stimuli (lines varying in length). In addition, we also sought to examine whether delayed retesting of either type of nontemporal stimuli would result in systematic distortion to the reference memory representation of the standard as observed in Ogden et al. (2008, 2010) using auditory temporal stimuli.

The results of Experiments 1 and 2 indicate that when multiple standards of different pitches or physical lengths are encoded into reference memory and maintained over a period of delay, generalization performance deteriorates, and correct identifications of the standard are reduced. Deterioration is greater when a more difficult comparison/standard ratio is employed. Under easy task conditions, interference appears greater in the auditory task than in the visual task, mirroring the findings of the auditory and visual temporal generalization experiments reported in Ogden et al. (2010). However, when task difficulty is increased (to a level comparable with the temporal research), interference effects appear comparable in both modalities, with significant increases in the variability of responding during delayed retesting on the pitch and line length generalization tasks. Regardless of the level of task difficulty or the modality of the standards when a nontemporal generalization task was employed, there was no evidence of the systematic distortions to the memory of the first standard characterized by a shift in the location of peak responding as observed in Ogden et al. (2008, 2010).

The absence of any systematic distortion to the memory of the first standard during visual temporal generalization (Ogden et al., 2010) or auditory and visual nontemporal generalization indicates that this type of interference may be unique to auditory temporal information. It is unclear why requiring participants to hold multiple standards over a period of delay would only lead to systematic interference when auditory temporal information is encoded, particularly given previous research that generally suggests that auditory temporal information is better remembered than visual temporal information. For example, rhythm memory has been found to be superior when the rhythm is presented in the auditory modality (Glenberg & Jona, 1991), and the timing of auditory stimuli is consistently found to be more accurate than that of visual stimuli (Wearden, Edwards, Fakhri, & Percival, 1998). Auditory stimuli are also more readily processed than visual stimuli (Jaskowski, Jaroszyk, & Hojan-Jezierska, 1990), and when auditory and visual stimuli are presented simultaneously, the auditory stimuli appear to contribute more to the timing judgement than do the visual stimuli (Penney, Gibbon, & Meck, 2000). One speculative possibility is that it is more difficult for participants to apply the rule "I can't remember the first standard but I remember that it was shorter/long higher/lower than the second standard" (which Ogden et al., 2008, 2010, suggested results in systematic interference to temporal stimuli) to nontemporal auditory or visual stimuli or to visual temporal stimuli. This points to the idea that there is something peculiar to auditory temporal representation, which has been hinted at in other studies; for example, it now appears that the classic "sounds are judged longer than lights" effect is driven by the way in which auditory stimuli are timed (see Jones, Poliakoff, & Wells, 2009, for a discussion of this issue).

Although systematic inference effects were not observed when multiple nontemporal standards were encoded, there were significant deteriorations in correct identifications of the standard during delayed retesting. It would therefore seem that rather than auditory temporal entries into reference memory being particularly vulnerable to interference, visual temporal entries appear to be particularly resistant to inference. It is unclear why visual temporal entries into reference memory would be more resistant to interference than auditory temporal, or visual or auditory nontemporal entries. One possible suggestion is that although auditory stimuli are more readily processed than visual stimuli (Jaskowski et al., 1990), they may be more vulnerable to overwriting. For example, Deutsch (1970) and Pechmann and Mohr (1992) have reported that pitches are particularly vulnerable to overwriting by other pitches. Deutsch and Pechmann and Mohr both used a discrimination task in which participants were presented with a standard pitch, followed by an interstimulus interval (ISI), followed by a comparison pitch. The task was for the participants to state whether the comparison was the same pitch as the standard. Deutsch and Pechmann and Mohr both found that the presentation of filler tones in the ISI led to a disruption in performance on the discrimination task, even when the filler tones were not attended to. Massaro (1970) also suggests that participants who attempt overt rehearsal of a standard pitch by humming the pitch during the ISI of a comparison task will perform worse than those who do not attempt overt rehearsal. Keller, Cowan, and Saults (1995), however, found that when covert rehearsal (imagining the pitch of the standard) was performed during the ISI, performance was superior to when an auditory or visual distractor task was performed in the ISI. It would therefore seem that merely hearing another pitch prior to retrieval of an existing pitch (or possibly auditory stimuli) could be sufficient to lead to interference in the memory of the previously encoded pitch. This apparent susceptibility to overwriting could also explain why there was a greater level of interference to pitch generalization performance than line length generalization performance under easy testing conditions.

CONCLUSIONS

Comparison of the results of Experiments 1 and 2 with those of Ogden et al. (2008, 2010) suggests that systematic interference to representations in reference memory only occurs when multiple auditory temporal stimuli are encoded and maintained. When nontemporal standards are employed, generalization performance becomes significantly more variable when multiple standards are encoded; however, this never results in systematic interference. It remains unclear why visual temporal information appears relatively resistant to interference when auditory temporal and auditory and visual nontemporal information appears to degrade as memory load and maintenance duration increase, or why only auditory temporal information is vulnerable to systematic interference.

These experiments touch on a more general and unexplored aspect to human memory. To date, the course of memory research in humans has focused on the encoding, retention, and retrieval of memory items, be they words, pictures, numbers, or episodic events. All of these types of item share a common characteristic in that they can be verbalized, and this verbalization can be the item that is stored. In fact, aside from the study of memory for faces, the issue of memory for stimulus attributes that cannot be so easily verbalized and encoded (e.g., texture, taste, length, volume, shape, smell, etc.) has seldom been explored. To be clear, the overall question is: How do we remember the texture, size, or weight of an object, or the taste of a food item? What characteristics does this type of memory exhibit? Is it open to higher cognitive manipulation or does it exist at a lower perceptual level, not with a verbal code but with a direct code? The only area of perception that has explored this issue in any degree of detail is the field of time perception. The results of this research have revealed some surprising and intriguing characteristics for this type of memory—that is, subjective shortening for short-term retention and systematic interference for multiple long-term items. In order to move this field forward, it is now necessary to explore this memory type across multiple modalities, to see the degree to which memory for stimulus attributes shows common characteristics (and, if so, then to map in detail what these characteristics are), or the degree to which this memory type is specific to a particular stimulus characteristic. We hope this current article represents a step along this road.

Original manuscript received 20 May 2009
Accepted revision received 26 October 2010
First published online 22 February 2011

REFERENCES

Conway, M. A., & Gathercole, S. E. (1987). Modality and long-term memory. *Journal of Memory and Language*, *39*, 371–391.

Deutsch, D. (1970). Tones and numbers: Specificity of interference in short-term memory. *Science*, *168*, 1604–1605.

Droit-Volet, S., Clément, A., & Wearden, J. H. (2001). Temporal generalization in 3-to 8-year-old children. *Journal of Experimental Child Psychology*, *80*, 271–288.

Elvevag, B., Brown, G. D. A., McCormack, T., Vousden, J. I., & Goldberg, T. E. (2004). Identification of tone duration, line length, and letter position: An experimental approach to timing and working memory deficits in schizophrenia. *Journal of Abnormal Psychology*, *113*, 509–521.

Engle, R. W, & Mobley, L. A. (1976). The modality effect: What happens in long-term memory? *Journal of Verbal Learning and Verbal Behavior*, *15*, 519–527.

Gibbons, H., Brandler, S., & Rammsayer, T. H. (2003). Dissociating aspects of temporal and frequency processing: A functional ERP study in humans. *Cortex*, *39*, 947–965.

Glenberg, A. M., & Jona, M. (1991). Temporal coding in rhythm tasks revealed by modality effects. *Memory and Cognition*, *5*, 514–522.

Harrington, D. L., Haaland, K. Y., & Knight, R. T. (1998). Cortical networks underlying mechanisms of time perception. *The Journal of Neuroscience*, *18*, 1085–1095.

Jaskowski, P., Jaroszyk, F., & Hojan-Jezierska, D. (1990). Temporal-order judgments and reaction time for stimuli of different modalities. *Psychological Research*, *52*, 35–38.

Jones, L. A., Poliakoff, E., & Wells, J. (2009). Good vibrations: Human interval timing in the vibrotactile modality. *Quarterly Journal of Experimental Psychology*, *62,* 2171–2186.

Keller, T. A., Cowan, N., & Saults, J. S. (1995). Can auditory memory for tone pitch be rehearsed? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21,* 635–645.

Kirsner, K. (1974). Modality differences in recognition memory for words and their attributes. *Journal of Experimental Psychology*, *103,* 579–584.

Lehman, E. B. (1982). Memory for modality: Evidence for an automatic process. *Memory & Cognition*, *10,* 554–564.

Maquet, P., Lejeune, H., Pouthas, V., Bonnet, M., Casini, L., Macar, F., et al. (1996). Brain activation induced by estimation of duration. *NeuroImage*, *3,* 119–126.

Massaro, D. W. (1970). Retroactive interference in short-term recognition memory for pitch. *Journal of Experimental Psychology*, *83,* 32–39.

McCormack, T., Brown, G. D. A., Maylor, E. A., Darby, R. J., & Green, D. (1999). Developmental changes in time estimation: Comparing childhood and old age. *Developmental Psychology*, *355,* 1143–1155.

McCormack, T., Brown, G. D. A., Maylor, E. A., Richardson, L. B., & Darby, R. J. (2002). Effects of aging on absolute identification of duration. *Psychology and Aging*, *17,* 363–378.

McCormack, T., Brown, G. D. A., Smith, M. C., & Brock, J. (2004). A timing-specific memory distortion effect in young children. *Journal of Experimental Child Psychology*, *8,* 33–56.

Murdock, B. B., & Walker, K. D. (1969). Modality effects in free recall. *Journal of Verbal Learning and Verbal Behavior*, *8,* 665–676.

Nairne, J. S., & Walters, V. L. (1983). Silent mouthing produces modality- and suffix-life effects. *Journal of Verbal Learning and Verbal Behavior*, *22,* 475–483.

Ogden, R. S., Wearden, J. H., & Jones, L. A. (2008). The remembrance of times past: Interference in temporal reference memory. *Journal of Experimental Psychology: Human Perception and Performance*, *34,* 1524–1544.

Ogden, R. S., Wearden, J. H., & Jones, L. A. (2010). Are memories for duration modality specific? *Quarterly Journal of Experimental Psychology*, *63,* 65–80.

Pechmann, T., & Mohr, G. (1992). Interference in memory for tonal pitch: Implications for a working-memory model. *Memory and Cognition*, *20,* 314–320.

Penney, T. B., Gibbon, J., & Meck, W. H. (2000). Differential effects of auditory and visual signals on clock speed and temporal memory. *Journal of Experimental Psychology: Human Perception and Performance*, *26,* 1770–1787.

Pouthas, V., Garnero, L., Ferrandez, A.-M., & Renault, B. (2000). ERPs and PET analysis of time perception: Spatial and temporal brain mapping during visual discrimination tasks. *Human Brain Mapping*, *10,* 49–60.

Wearden, J. H. (1992). Temporal generalization in humans. *Journal of Experimental Psychology: Animal Behavior Processes*, *18,* 134–144.

Wearden, J. H. (1999). "Beyond the fields we know . . . ": Exploring and developing scalar timing theory. *Behavioural Processes*, *45,* 3–21.

Wearden, J. H., Edwards, H., Fakhri, M., & Percival, A. (1998). Why "sounds are judged longer than lights": Application of a model of the internal clock in humans. *The Quarterly Journal of Experimental Psychology*, *51B,* 97–120.

Wearden, J. H., & Ferrara, A. (1993). Subjective shortening in humans' memory for stimulus duration. *Quarterly Journal of Experimental Psychology*, *46B,* 163–186.

Wearden, J. H., Parry, A., & Stamp, L. (2002). Is subjective shortening in human memory unique to time representations? *Quarterly Journal of Experimental Psychology*, *55B,* 1–25.