

## Short-Term Forgetting Without Interference

Denis McKeown  
University of Leeds

Tom Mercer  
University of Wolverhampton

In the 1st reported experiment, we demonstrate that auditory memory is robust over extended retention intervals (RIs) when listeners compare the timbre of complex tones, even when active or verbal rehearsal is difficult or impossible. Thus, our tones have an abstract timbre that resists verbal labeling, they differ across trials so that no “standard” comparison stimulus is built up, and the spectral change to be discriminated is very slight and therefore does not shift stimuli across verbal categories. Nonetheless, performance in this nonverbal immediate memory task was better at short (1-, 2-, or 4-s) than long (8-, 16-, or 32-s) RIs, an outcome predicted by temporal distinctiveness theory whereby at long RIs, tones are closer in time to tones on previous trials. We reject this account in the 2nd experiment, where we demonstrate that the ratio of RI to intertrial interval makes absolutely no difference to performance. We suggest that steady forgetting is consistent with a psychoacoustically derived conception of an auditory memory (the *timbre memory model*) that embodies time-based forgetting in the absence of feature-specific interference.

*Keywords:* auditory memory, temporal distinctiveness, sensory memory

The preoccupations and frustrations of the earliest memory researchers (e.g., Angell & Harwood, 1899; Whipple, 1901) remain to the fore in the theoretical conceptions of immediate memory today. Notably, despite more than a century of investigation, there is still disagreement over the role that time plays in forgetting over the short-term (e.g., Barrouillet, Portrat, Vergauwe, Diependaele, & Camos, 2011; Oberauer & Lewandowsky, 2008). Researchers have actually proposed two distinct forgetting functions, which ascribe some role to time—passive trace decay and temporal distinctiveness. Decay suggests that a memory trace simply deteriorates and fades away over the passage of time, without any particular causal agent. Information loss is therefore believed to be a consequence of the total (or *absolute*) amount of time that has elapsed since the presentation of an event (Cowan, Saults, & Nugent, 1997). Conversely, temporal distinctiveness posits that the probability of recalling an item depends on how unique that item is within a particular temporal context (Oberauer & Lewandowsky, 2008). That is, temporally isolated items should be more likely to be recalled than temporally crowded items (Lewandowsky, Nimmo, & Brown, 2008); hence, distinctiveness attributes an important role to the passage of *relative* time (or the temporal separation of events in memory) rather than absolute time (Cowan et al., 1997). According to this view, forgetting is a form of time-based proactive interference, and distinctiveness models—such as G. D. A. Brown, Neath, and Chater’s (2007) SIMPLE (Scale-Independent Memory, Perception, and LEarning) model—reject the notion of trace decay.

Other theoretical accounts of short-term memory have also abandoned decay in favor of interference (e.g., the serial-order-in-a-box model of Farrell & Lewandowsky, 2002), and Nairne (2002) influentially noted that the idea of decay, and the very notion of an activated trace in a short-term store, ignores the influence of the environment in which a stimulus occurs and the events surrounding the to-be-remembered stimulus. Even so, actually testing whether decay plays a role in information loss is very challenging. Specifically, if there is no evidence of forgetting over time—contrary to the expectations of decay—it is possible for theorists to argue that auxiliary processes such as rehearsal were in operation (Lewandowsky, Oberauer, & Brown, 2009). Convincing evidence against decay must therefore demonstrate the persistence of memory over time while preventing compensatory rehearsal, a “core difficulty in the study of time and memory” (Lewandowsky, Geiger, Oberauer, & Morrell, 2010, p. 959). One solution is to insert some form of secondary task into the procedure that obstructs rehearsal, but this carries the risk of inadvertently introducing interference that may mimic time-based forgetting (see Dewar, Cowan, & Della Salla, 2007, for a demonstration of nonspecific interference caused by a variety of distractor tasks). Researchers of verbal memory have recently designed some ingenious ways of obstructing rehearsal while preventing interference (e.g., Berman, Jonides, & Lewis, 2009; T. Green & McKeown, 2007; Lewandowsky, Duncan, & Brown, 2004; Oberauer & Lewandowsky, 2008), but an alternative approach is to rely on procedures that require retention of *nonverbal* stimuli.

By their nature, these stimuli cannot be verbally imitated, and this allows the impact of absolute time to be explored without the problem of compensatory rehearsal. One recent and excellent example of this approach was highlighted by Ricker and Cowan (2010) in a study of visual working memory. On each trial, they showed participants three unconventional symbols in different positions for 750 ms, and in an earlier experiment they had confirmed that these stimuli were not verbally encoded (Ricker,

---

This article was published Online First March 26, 2012.

Denis McKeown, Institute of Psychological Sciences, University of Leeds, Leeds, England; Tom Mercer, School of Applied Sciences, University of Wolverhampton, Wolverhampton, England.

Correspondence concerning this article should be addressed to Denis McKeown, Institute of Psychological Sciences, University of Leeds, Leeds LS2 9JT, United Kingdom. E-mail: d.mckeown@leeds.ac.uk

Cowan, & Morey, 2010). The visual array was followed by a mask (100 ms) and then a retention interval (RI) lasting for 1, 3, or 6 s. In the recognition phase, another symbol was presented, and participants had to decide whether they had seen that character in that position at the start of the trial. Crucially, Ricker and Cowan found a significant effect of the passage of time, with accuracy at 1.5 s being higher than at 3 s and 6 s. Performance at 3 s was also better than 6 s. This pertinent finding seemed to show the detrimental effect of the passage of absolute time, but Ricker and Cowan have not ruled out a potential role for distinctiveness. There is one other possible explanation too. Ricker and Cowan analyzed the proportion of hits (correctly detecting a difference) and correct rejections (accurately identifying that stimuli were the same) in addition to overall task accuracy. They reported no effect of the duration of the interval for hits, but there was a significant drop in correct rejections over time. Perhaps, then, their time-based performance decline was due to a shift in criterion—that is, at longer gaps, participants may have changed their decision criterion, being more likely to state that stimuli were different (a change in response bias). The memory trace may not have suffered any loss in fidelity, but a form of criterion noise may have been added to the representation. Despite this possibility, numerous experiments using auditory stimuli have found that the absolute amount of time separating two tones is important. The principle of these auditory studies was similar to Ricker and Cowan: They utilized stimuli that differed subtly from one another (typically in pitch) and were not susceptible to verbal encoding. On each trial, participants compared two tones over a silent RI, deciding whether they were the same or different, and usually task accuracy became gradually worse as the interval between the sounds was extended (e.g., Clément, Demany, & Semal, 1999; Demany, Trost, Serman, & Semal, 2008; Harris, 1952; Kaernbach, 2004; Kaernbach & Schlemmer, 2008; Keller, Cowan, & Saults, 1995; Kinchla & Smyzer, 1967; Moss, Myers, & Filmore, 1970). Supporting this behavioral literature, neurophysiological evidence has placed the lifetime of auditory memory in the region of about 10 s (Sams, Hari, Rif, & Knuutila, 1993), and a recent model of sound recognition hypothesized a key role for decay (McLachlan & Wilson, 2010).

However, Cowan et al. (1997) outlined a major flaw in the majority of previous studies exploring decay in nonverbal auditory memory. Most of the investigations using the two-tone comparison procedure had extended RIs separating the tones on a trial but had very brief gaps between trials. As a result, there may have been a loss of temporal distinctiveness at longer RIs, and, consequently, tones from *previous* trials may have influenced and distorted the representation on the *current* trial. Proactive interference may therefore have been responsible for the drop in performance. To test this idea directly, Cowan et al. (1997) examined the impact of both absolute time (the total amount of time that had elapsed since presentation of the to-be-remembered tone) and relative time (the temporal separation of tones on the current trial from those in the recent acoustical past). Their participants compared pure tones varying in frequency over RIs of 1.5, 3, 6, and 12 s (absolute time), and the intertrial interval (ITI) was varied between 3, 6, 12, and 24 s (relative time). A role for relative time was observed, but absolute time appeared to be important too. The latter finding was compatible with trace decay, but Cowan, Saults, and Nugent (2001) reexamined their 1997 data and assessed the influence of

RIs and ITIs from a number of previous trials. Impressively, they found no evidence for time-based forgetting on trials where the tones on the current trial were maximally separated from previous sounds—regardless of the RI. This striking finding strongly favors distinctiveness over trace decay, but unfortunately Cowan et al. (2001) did not conduct any inferential statistics because they were unable to rely on their full 1997 data set (they had to pool the results, and the contribution of the participants differed across the various RIs). Hence, their findings appear inconclusive.

In summary, a pertinent question still awaits an answer: Is there any evidence for time-based forgetting in nonverbal short-term memory? The purpose of the present study was to assess the influence of both absolute (trace decay) and relative (distinctiveness) time in memory for timbre. *Timbre* is a complex, multidimensional attribute usually defined as the “quality” that makes a sound unique (Marozeau, de Cheveigné, McAdams, & Winsberg, 2003; Menon et al., 2002), and it is a particularly useful attribute to study because it cannot be verbally encoded—at least not in an accurate or reliable manner. Crowder (1989) noted that “humans are utterly incapable of reproducing physically any but the grossest dynamic or spectral features of timbre” (p. 478), and evidence supporting the nonverbal encoding of the timbre of complex tones was recently reported by McKeown, Mills, and Mercer (2011). Listeners compared subtle timbre differences between two complex tones separated by an extended RI, and in one condition, they read aloud during that interval. The rationale was simple: If the memoranda were being verbally categorized or labeled, the reading task would be expected to hinder this, significantly decreasing discriminatory performance in comparison with a silent RI condition. Contrary to this prediction, there were no differences between silent and reading conditions, supporting the notion that the memory for the timbre of complex tones is encoded as an abstract, nonverbal representation.

The above demonstration of the nonverbal memory for complex sounds—and other recent findings in our laboratory on the effects of pretrial cues on timbre discriminations (McKeown & Wellsted, 2009) and influences of interfering distractor tones on memory for timbre (Mercer & McKeown, 2010a, 2010b)—has supported our conception of an auditory memory (the timbre memory model [TMM]) that serves to track recent auditory stimulation in the form of a spectro-temporal code. This is an auditory memory for timbre that conforms to a number of the criteria traditionally thought to define a short sensory store, such as being independent of attention, of very fine resolution, and specific to a single modality (Cowan, 2008; Winkler & Cowan, 2005), but its persistence over tens of seconds (Mercer & McKeown, 2010b) suggests that that traditional definition may not apply (see also Demany & Semal, 2008, for a discussion of the properties of short-term auditory memory). By maintaining a record of the immediate past, TMM also serves as a *predictor* of the running changes in auditory stimulation (by comparing what has recently occurred with current stimulation, TMM is able to identify changes and update itself). This record of the recent auditory past is conceived as a footprint or residue within frequency- or channel-specific adapted populations of fibers (see, e.g., May & Tiitinen, 2010, for a description of stimulus-specific adaptation). Therefore adapted or suppressed populations of neurons representing information in distinct frequencies or channels constitute the memory of the system, a record that we propose does not demand rehearsal or attention for main-

tenance. Forgetting is envisaged as a slow recovery from adaptation in previously activated channels (McKeown et al., 2011) or as a result of interference between successive sound “features” that share frequency channels (Mercer & McKeown, 2010b). Thus, we favor *both* time-based forgetting and feature-specific interference. The present experiments address the former while controlling for the latter.

### Experiment 1

Experiment 1 aimed to assess whether discriminatory performance was affected by the passage of absolute time in a two-tone comparison procedure. On each trial, listeners heard a standard tone followed by a silent RI lasting for 1, 2, 4, 8, 16, or 32 s.<sup>1</sup> A comparison tone was then presented, and participants had to decide whether it differed from the standard. There were a number of crucial controls in place to ensure that the tones were nonverbal in nature. Firstly, the discrimination was based on very subtle changes to the spectra of the tones that prevented accurate categorical or verbal translation (see Mercer & McKeown, 2010b, for a full discussion of this issue). Specifically, complex sounds were generated with a slight increment to four of their components. This increment was either 5.1 or 3.5 dB sound-pressure level (SPL), allowing the role of task difficulty to be examined. Listeners were also prevented from building up a session long “standard signal” representation by roving the pitch of the tones between trials. This ensured that the sounds did not lend themselves to verbal encoding of a form, such as “signal-like” or “non-signal-like,” and prevented participants from comparing tones with a long-term standard memory (for a theoretical account, see Bull & Cuddy, 1972; Durlach & Braida, 1969; Durlach, Braida, & Ito, 1986). Hence, sounds were abstract, nonverbal, and non-category-like. In addition, effort was made to reduce or eliminate proactive interference by separating the sound pairings by extended ITIs of 34 s (plus response time). This should diminish the influence of sounds from the recent acoustical past and prevent a loss of distinctiveness at longer gaps (because the ITI was always lengthier than the RI). Roving pitch between trials was designed to reduce proactive interference too. Mercer and McKeown (2010b) varied tone pitch across trials and demonstrated that proactive interference was heightened when a prior interfering tone was spectrally similar to the current trial stimuli; conversely, proactive interference was absent when the prior tone differed spectrally from the trial tone (the pitch of a complex harmonic tone determines its spectral content). Varying the sounds spectrally across trials should thus remove proactive interference from old trials. Given this, will the passage of absolute time still lead to forgetting?

### Method

**Participants.** Five individuals (three men; mean age = 29.6 years, range = 21–54) with self-reported normal hearing participated in the experiment. This included the two authors, two individuals with prior experience of psychoacoustical memory tasks (identified as KL and NS), and one inexperienced listener (identified as FB). FB, KL, and NS received hourly payment for participation.

**Stimuli and materials.** Six periodic complex tones consisting of eight frequencies were generated for this experiment and are displayed in Figure 1. Four components within each complex tone

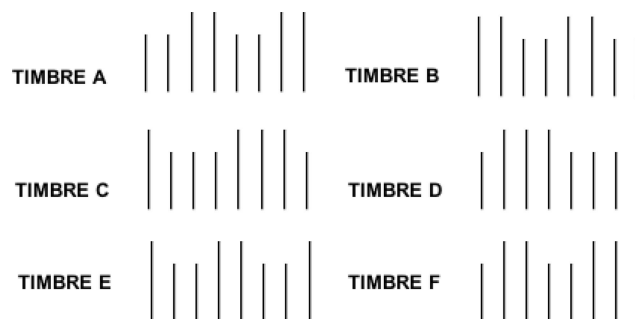


Figure 1. Diagrammatic representation of the six complex tones, pairings of which were used as the standard and comparison tones in Experiment 1. Each line represents a single harmonic, increasing in frequency left to right. The height of each line indicates appropriate *relative* intensities of the harmonics. The longer lines reflect the frequency components that were incremented in intensity by 5.1 dB or 3.5 dB.

were incremented by an additional 5.1 or 3.5 dB SPL (taller lines in Figure 1), and this manipulation altered the spectral profile of the stimuli. The 5.1-dB condition represents a slight increase in feature prominence within the spectral profile of stimuli relative to the 3.5-dB condition. All tones were 300 ms in duration, including a 10-ms cosine onset and offset ramp, and were approximately 80 dB in intensity (Brüel & Kjør type 2260H). The tones could be presented at any of five different pitches between trials (from D at 146.8 Hz through F#, B, G1, and Bb1 at 466.2 Hz).

Stimuli were generated using Tucker–Davis Technologies (TDT) RP2.1 hardware and a PC running TDT RPvdsEx and Mathworks MATLAB software. A custom designed MATLAB program also presented trials and recorded participant responses. Stimuli were attenuated (TDT PA5), filtered (100 Hz to 10 kHz, Kemo VBF21M), and output to the left earpiece of STAX SR-303 Classic headphones via a STAX SRM-323 II unit and TDT HB7 headphone amplifier. Participants completed the study seated within an Industrial Acoustics Company double-walled sound-attenuating booth.

**Design and procedure.** Each trial began with the presentation of the standard tone followed by the comparison tone after a silent RI. This interval was varied in different conditions between 1, 2, 4, 8, 16, and 32 s. Participants had to decide whether the two tones were the same or different in terms of timbre, indicating their answer on a response box without feedback. The next trial began 34 s after a response was made, and this extended ITI ensured that the standard tone was always temporally closer to the comparison on the *current* trial rather than to the comparison on the *previous* trial.

On each trial, a complex tone was presented either with itself (“same” trials; AA, BB, CC, DD, EE, FF) or a single other tone (“different” trials; AB, BA, CD, DC, EF, FE). This ensured that

<sup>1</sup> In recent studies of time-based forgetting in visual memory, a mask has often been employed to eliminate sensory memory. Although such a mask was considered in the design of this study, our concern was that the use of such a mask with auditory stimuli would produce some form of interference or backward masking. Because the primary purpose of the experiment was to assess whether there was time-based forgetting, we therefore preferred to ensure we did not introduce any external interference in the task.

when stimuli differed, the difficulty of the discrimination was similar regardless of the trial type.<sup>2</sup> As such, on “different” trials, the comparison tone was an inverse of the standard (see Figure 1). The use of six different standard tones and the roving pitch arrangement ensured that the listeners were not able to build up a long-term memory of the sounds or rely on a form of context-coding (Durlach & Braida, 1969). The tones were abstract and non-category-like.

All participants began the experiment comparing tones with the distinguishing components raised by 5.1 dB (i.e., the easier discrimination). These components were then reduced to a 3.5-dB intensity increment for the second phase of the study. This arrangement manipulated the subtlety of the discrimination and thereby allowed the effects of feature prominence to be examined. Although no extensive training phase was incorporated into this design, a practice session preceded both the 5.1-dB and 3.5-dB conditions and required participants to contrast standard and comparison tones over a 1-s RI. This was intended to familiarize individuals with the stimuli and the task. Following this, the experimental sessions commenced. Each of these sessions lasted for approximately 30 min and involved two trial blocks of a randomly determined RI (e.g., 4 s). Each experimental block contained an equal number of “same” and “different” trial types, although the exact amount of trials was reduced with longer RIs to keep each session equivalent in duration (e.g., 20 trials for the 1-s interval, 12 trials for the 32-s interval). Overall, listeners were presented with 30 “same” and 30 “different” trial types for each of the six RIs, and the 12 major trial types were presented once for every pitch. Participants completed 360 trials for each intensity increment condition and completed 720 experimental trials in total. This required approximately 5.5 hr of testing completed in individual sessions three to four times a week.

## Results and Discussion

Performance at each RI was indexed by  $d'$ —a bias-free measure of sensitivity that is not affected by changes in criterion. To calculate this,  $P_{ss}$  (the probability of responding “same” when stimuli were same) and  $P_{sd}$  (the probability of responding “same” when stimuli were different) were computed and converted into  $d'$  (see Bi, 2002, for a full explanation of calculating  $d'$  in the same-different task). Cases of extreme performance (i.e., 100%  $P_{ss}$  rate or 0%  $P_{sd}$  rate) were corrected using the log-linear rule. This strategy entails adding 0.5 to all  $P_{ss}$  and  $P_{sd}$  values while increasing the overall trial number to  $N + 1$  (e.g., Snodgrass & Corwin, 1988). This is deemed to be an acceptable method of dealing with extreme sensitivity (G. S. Brown & White, 2005), and such a correction was necessary for four of the participants (DM, FB, KL, and NS). However, only nine cases of extreme performance were observed throughout the entire experiment.

The  $d'$  was computed for the group (based on  $P_{ss}$  and  $P_{sd}$  values averaged across all participants) and for each individual. These data are shown in the six panels of Figure 2. Average performance for all participants, shown in Panel 1, revealed a steady decrease in task accuracy as the RI was extended, and this effect was more pronounced for the 3.5-dB conditions. However, this pattern is more complex when considering individual participant data. Discriminatory accuracy was usually better at shorter RIs, but there were some cases where  $d'$  was actually larger at longer RIs. It

should also be noted that, overall, performance on the task was remarkably impressive, even after very long intervals.

### Examining the Influence of Absolute Time

To assess the effect of RI duration, Marascuilo’s (1970)  $K$ -signal significance test was conducted. The  $K$ -signal test statistic allows multiple  $d'$  values to be compared and resembles the chi-square test in that observed and expected  $d'$  values are contrasted. The resulting  $K$  value has an asymptotic chi-square distribution with  $K - 1$  degrees of freedom (Marascuilo, 1970) and indicates whether there is a significant main effect. The  $K$ -signal test was run separately for the 5.1-dB and 3.5-dB conditions. A significant main effect of time was found in the 5.1-dB condition for the group,  $\chi^2(5) = 13.11, p = .022$ , and for participants DM,  $\chi^2(5) = 19.61, p = .002$ ; KL,  $\chi^2(5) = 13.01, p = .023$ ; NS,  $\chi^2(5) = 11.28, p = .046$ ; and TM,  $\chi^2(5) = 14.27, p = .014$ . As recommended by Bi, Ennis, and O’Mahony (1997), these significant main effects were followed-up with an algorithm described by Gourevitch and Galanter (1967). This statistic resembles a  $Z$  test in that the difference between two  $d'$  scores is calculated and divided by the square root of the summed variance for both  $d'$  values. The resulting outcome is approximately normally distributed and can then be assessed for significance in a normal distribution table. The results of this test are shown in Table 1. Although there were some differences between individuals,  $d'$  at 32 s was always significantly lower than at 1 s, and—excluding NS—performance at 1 s was significantly higher than 16 s too. Additionally,  $d'$  at 2 s was significantly higher than at 32 s, with the exception of KL. Conversely, participant FB did not show a main effect of time,  $\chi^2(5) = 5.1, p = .404$ .

A similar pattern emerged for the 3.5-dB conditions. The  $K$ -signal test found significant main effects of time for the group,  $\chi^2(5) = 13.09, p = .023$ , and for participants DM,  $\chi^2(5) = 23.09, p < .001$ ; KL,  $\chi^2(5) = 20.25, p = .001$ ; NS,  $\chi^2(5) = 11.44, p = .043$ ; and TM,  $\chi^2(5) = 23.38, p < .001$ . Gourevitch and Galanter’s (1967) post hoc test (also shown in Table 1) typically documented a more dramatic decline in sensitivity as the RI was extended beyond 1 s. Indeed, participants KL and TM both performed significantly worse at every RI in comparison with 1 s. Conversely, FB again showed no effect of time,  $\chi^2(5) = 6.34, p = .275$ , although there was a modest decline in task accuracy.

### Examining the Influence of Feature Prominence

The  $d'$  values in the 5.1-dB and 3.5-dB conditions were contrasted at each RI using Gourevitch and Galanter’s (1967) statistic. This test highlighted that sensitivity in the two intensity conditions rarely differed, suggesting little impact of these relatively small differences in feature prominence. However, to increase the power of this analysis,  $P_{ss}$  and  $P_{sd}$  scores were pooled across RIs and were converted into an overall  $d'$  for 5.1 dB and 3.5 dB. The results, displayed in Figure 3, indicate a modest decline in sensitivity when features were less prominent. An analysis employing

<sup>2</sup> It was found that the pairing of certain complex tones was extremely easy, whereas the pairings of others was virtually indistinguishable. The three major pairings (A and B, C and D, E and F) were chosen to be reasonably similar in terms of difficulty.



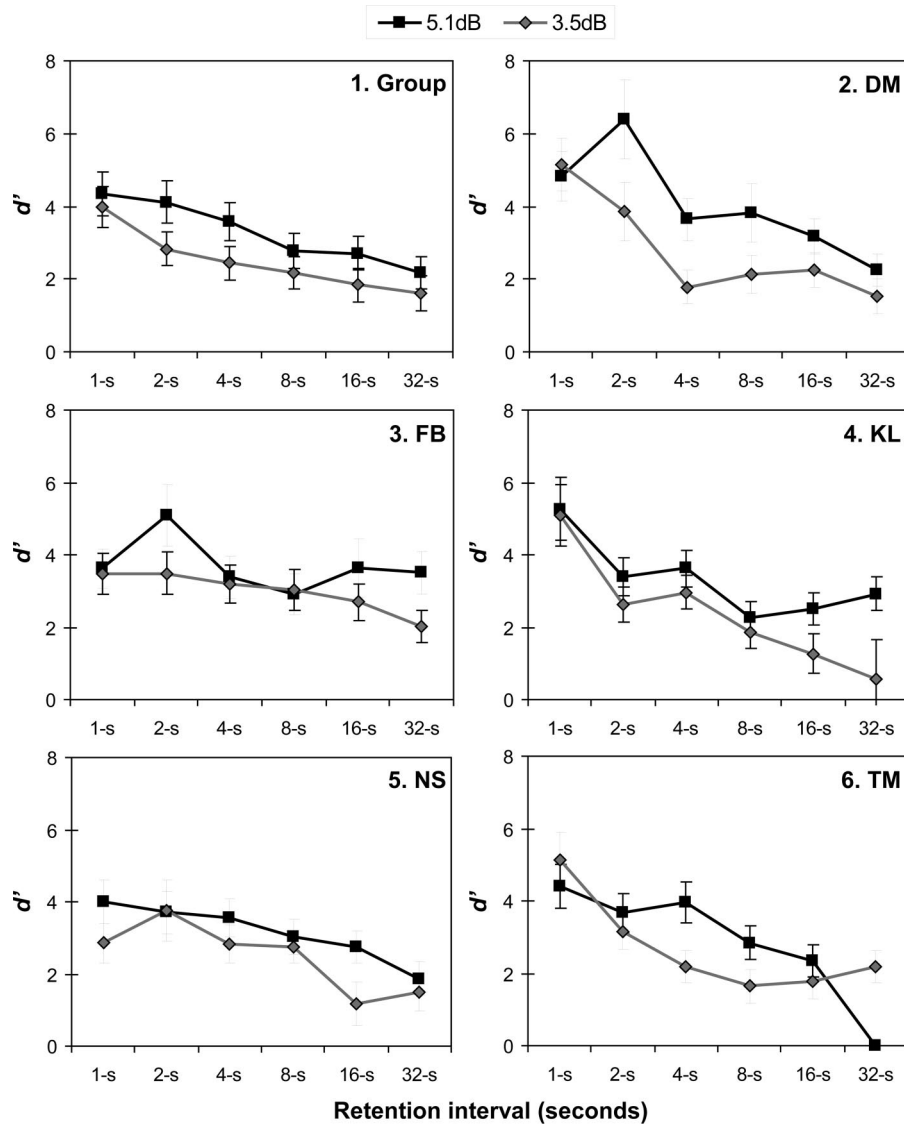


Figure 2. Task performance in Experiment 1 for each retention interval. Accuracy is expressed as  $d'$  for stimuli with 5.1-dB (black lines) and 3.5-dB (gray lines) increments. Panel 1 shows performance for the entire group (collapsed  $d'$ ), and the remaining five panels show individual participant data. Error bars reflect  $\pm 1$  SE of  $d'$  (see Macmillan & Creelman, 2005). No error bar is shown for TM in the 5.1-dB, 32-s condition because performance was at chance. Variance of  $d'$  had to be estimated in the analyses for this participant.

Gourevitch and Galanter's test revealed that performance was significantly better at 5.1 dB than at 3.5 dB for the group ( $Z = 2.76, p = .006$ ) and for participants DM ( $Z = 3.01, p = .003$ ), FB ( $Z = 2.21, p = .027$ ), KL ( $Z = 3.53, p < .001$ ), and NS ( $Z = 2.71, p = .007$ ). Conversely, there was no difference for participant TM ( $Z = 1.04, p = .298$ ). A more sensitive analysis therefore revealed a significant effect of feature prominence in five of six cases. It appears that a tone with less prominent features is more likely to be lost.

In a final analysis,  $P_{ss}$  and  $P_{sd}$  were collapsed across 5.1- and 3.5-dB conditions. This allowed us to examine the impact of time regardless of feature prominence, and the averaged data for the group are shown in Figure 4. The  $K$ -signal test revealed a signif-

icant main effect of RI,  $\chi^2(5) = 25.21, p < .001$ , and these data clearly display the steady falling away of performance across the log-spaced time axis as an almost exact linear function.

In summary, in this nonverbal immediate memory (*nime*) task, we have jumped or avoided many of the hurdles or difficulties presented in tone comparison studies but still observe steady forgetting over time. Notably, we have used stimuli varying across a fairly wide range in fundamental frequency ( $f_0$ ) so that our listeners have been prevented from building up long-term "standards" throughout our experimental sessions, so as to tap the kind of more transient auditory memory for the standard stimulus that is our focus here. Furthermore, we have introduced quite subtle spectral features to the spectra of our tones so that when two

Table 1  
Comparison of the Six Conditions Using Gourevitch and Galanter's (1967) Test

Comparison	Group		DM		KL		NS		TM	
	5.1 dB	3.5 dB	5.1 dB	3.5 dB	5.1 dB	3.5 dB	5.1 dB	3.5 dB	5.1 dB	3.5 dB
1 s vs. 2 s	0.26	1.56	-1.24	1.20	1.84	2.52*	0.36	-0.91	0.91	2.23*
1 s vs. 4 s	0.93	2.13*	1.32	3.92****	1.63	2.18*	0.58	0.03	0.52	3.37****
1 s vs. 8 s	1.99*	2.48*	0.94	3.39****	3.10****	3.34****	1.27	0.15	2.02*	3.94****
1 s vs. 16 s	2.13*	2.91****	1.99*	3.34****	2.84****	3.78****	1.65	2.07*	2.67**	3.78****
1 s vs. 32 s	2.84****	3.21****	3.17****	4.15****	2.39*	3.28****	2.78**	1.79	2.56*	3.37****
2 s vs. 4 s	0.67	0.61	2.24*	2.25*	-0.30	-0.51	0.20	0.93	-0.38	1.45
2 s vs. 8 s	1.73	1.02	1.91	1.82	1.67	1.17	0.88	1.05	1.17	2.21*
2 s vs. 16 s	1.87	1.50	2.73**	1.73	1.30	1.89	1.27	2.49*	1.88	2.02*
2 s vs. 32 s	2.59**	1.84	3.55****	2.49*	0.69	1.75	2.41*	2.26*	2.17*	1.45
4 s vs. 8 s	1.10	0.42	-0.20	-0.50	2.02*	1.70	0.73	0.11	1.55	0.83
4 s vs. 16 s	1.23	0.92	0.62	-0.69	1.65	2.38*	1.14	2.04*	2.18*	0.64
4 s vs. 32 s	2.00*	1.27	1.88	0.37	1.02	2.07*	2.38*	1.75	2.32*	0.00
8 s vs. 16 s	0.12	0.85	0.71	-0.17	-0.40	0.86	0.42	2.07*	0.78	-0.17
8 s vs. 32 s	0.92	0.50	1.73	0.83	-1.04	1.12	1.72	1.77	1.70	-0.83
16 s vs. 32 s	0.82	0.36	1.39	1.04	-0.65	0.58	1.34	-0.40	1.39	-0.64

Note. The individual values denote Z scores. FB's data were not subjected to post hoc analyses because there was no main effect of time for this participant.

\*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .005$ . \*\*\*\*  $p < .001$ .

sounds differ on a trial, they do so well within a verbal category (i.e., the two complex tones on a trial do not extend across different verbal categories). Therefore, we believe we are investigating nonverbally coded memory, whereas most memory research cannot make that claim. Our approach is very much based on these two considerations, ones largely overlooked in prior research. However, as described above, our systematic manipulations of stimulus similarity in prior studies (Mercer & McKeown, 2010a, 2010b) also support our claim for avoiding one other major hurdle: proactive interference. To escape proactive interference, we have both changed sounds on successive trials to minimize spectral overlap (because similarity has been shown to increase proactive interference; cf. Visscher, Kahana, & Sekuler, 2009), and we have used very extended ITIs (believed to reduce proactive interference; cf. Cowan et al., 1997). Also, because we are not forced to circumvent possible verbal rehearsal in the RIs of our experiments,

we do not have to enter into the debate of whether the key manipulation to prevent such rehearsal—articulatory suppression (typically repeating an irrelevant word)—itself introduces another form of interference: retroactive interference by the verbal-rehearsal preventing activity. In the next experiment, we turn to a fresh difficulty not unrelated to these forms of stimulus proactive and retroactive interference: the possible role of stimulus distinctiveness or, more properly, local stimulus temporal distinctiveness.

### Experiment 2

In a short article entitled “The Importance of Temporal Distinctiveness for Forgetting Over the Short Term,” Unsworth, Heitz, and Parks (2008) have both summarized some key studies and presented a quite convincing set of data for the role of distinctiveness in forgetting, using “single” memoranda (as opposed to series

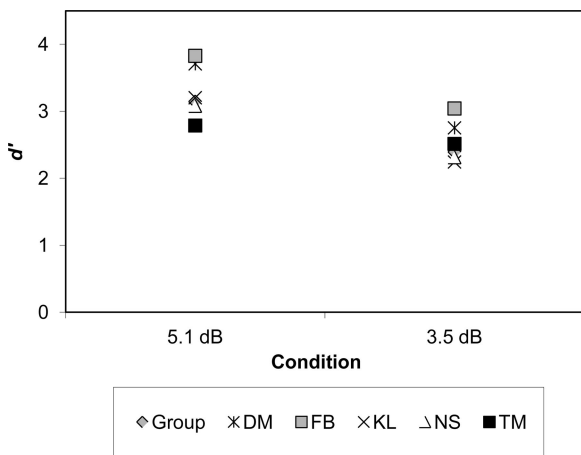


Figure 3. Data pooled across retention intervals for 5.1-dB and 3.5-dB conditions.

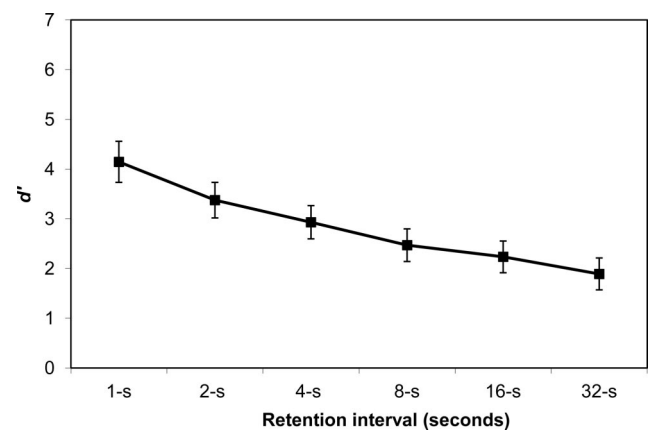


Figure 4. Group  $d'$  values in Experiment 1 for each retention interval averaged across 5.1- and 3.5-dB feature prominence conditions. Error bars reflect  $\pm 1$  SE of  $d'$ .

of successive stimuli and, therefore, more immediately relevant to our own situation). *Distinctiveness* here means that we may view stimuli as points retreating into the psychological past, and best remembered stimuli are predicted to be those that hold their distance (and, hence, autonomy from interference or confusability) from their neighbors. In a mixed design, Unsworth et al. tested recall of three-consonant trigrams (such as “VKR”) over varying RIs of 4, 8, 12, or 16 s (with a distractor task to prevent rehearsal). In the control condition, one group ( $N = 30$ ) experienced a fixed ITI of 1.5 s for all four RIs. In the next group ( $N = 28$ ), one only of the four RIs (that at 16 s) was preceded by a much longer ITI of 60 s. Finally, in the third group ( $N = 29$ ), one only of the four RIs (that at 8 s) was preceded by the ITI of 60 s. This arrangement nicely equated overall recall time across the groups in the experiment. In strong support for distinctiveness, performance in a long-ITI condition (with a RI of 16 s) showed *better* performance than a short-ITI condition (but with shorter RIs of 8 and 12 s). Increasingly distinct verbally coded items (those that followed extended ITIs) were best recalled—but what about nonverbal immediate recall?

As noted in the introduction to Experiment 1 above, in their tone-comparison study, Cowan et al. (1997) varied both absolute time in the RI and ITI, but Cowan et al. (2001) recognized that they had failed to consider the importance of “global” distinctiveness in their earlier study; that is, they did not control how distinct a particular item was from all other previously presented items. Fortunately Cowan et al. (1997) had recorded not only the current trial’s RI (trial  $N$ ) and preceding ITI (trial  $N - 1$ ) but also the RI from the previous trial (trial  $N - 1$ ) and the ITI prior to that. Figure 5 highlights this rather complicated arrangement. Because Cowan et al. (1997) had recorded the RI for the current trial ( $N$ ) and the preceding ITI, and the RI of the previous trial ( $N - 1$ ) and the ITI preceding that, Cowan et al. (2001) were able to reexplore distinctiveness in greater depth. They found that performance on standard-comparison pairs, which were maximally separated in time from previous pairs, were discriminated to a high level regardless of the RI separating the tones. Specifically, when the ITI from the previous two trials was 24 s and the RI separating the standard and comparison on trial  $N - 1$  was 12 s (i.e., the maximal intervals in all cases), they observed no information loss over time and suggested that, in conditions of maximum distinctiveness, there may be no temporal decay.

Unfortunately, any conclusion from this study must be taken with caution. To assess forgetting over time, Cowan et al. (2001) only contrasted performance with the shortest (1.5-s) and longest (12-s) RIs; however, because performance at 1.5 s was unusually poor (in comparison with other conditions), this may have inadvertently given the impression that there was no information loss in their condition of maximal distinctiveness. Furthermore, con-

trasting proportion of correct responses at 3 and 12 s in this condition *does* seem to indicate decline over time. The difficulty with Cowan et al.’s results is that, because it was a post hoc reexamination of their 1997 study, they could only rely on certain aspects of their results because particular data points were only available for some of the participants, and in many cases, there was only one instance of a particular trial of interest. Hence, these data do not yield strong evidence for distinctiveness effects.

What is needed clearly is a study that systematically varies intertrial and RIs for tone memory. Therefore, in our second experiment, listeners compared standard and comparison tones over intervals of 2 or 32 s and decided whether they were the same or different; however, to manipulate distinctiveness, the ITI in different conditions was set at 2, 16, or 34 s. This ensured that there was temporal crowding at the shortest ITI, but this became less prominent as the ITI was extended. Accounts based on distinctiveness would thus expect poorest performance with a 2-s ITI because this condition features the greatest degree of temporal crowding, and the best performance was hypothesized for the 34-s ITI. On the basis of our earlier research (Mercer & McKeown, 2010b), we would only anticipate interference from past trials when there was a high degree of feature overlap from past sounds in relation to the current tone. However, because in the present experiment pitch was varied on each trial ensuring that there was little spectral feature overlap between trials, such feature-based proactive interference should be minimal, and therefore we do not anticipate any significant effect of the ITI manipulation.

## Method

**Participants.** Five individuals (three men; mean age = 28.8 years, range = 19–56) with self-reported normal hearing participated in this experiment. This included the two authors and three individuals with no prior experience of psychoacoustical memory tasks (identified as MJ, RS, and VG).

**Stimuli and apparatus.** The six complex tones from the previous experiment were used in the present study. All other arrangements and apparatus were identical to Experiment 1, except the four incremented components were always increased by 5.1 dB.

**Design and procedure.** On each trial, participants compared standard and comparison complex tones over a silent RI and judged whether they differed in terms of timbre. The RI was either 2 or 32 s in different blocks, and the next trial began 2, 16, or 34 s after the participant responded. This yielded six possible conditions (RI:ITI): 2:2, 2:16, 2:34, 32:2, 32:16, and 32:34. All listeners began the experiment by comparing tones over a 1-s interval (to familiarize them with the stimuli and task) before proceeding to testing. Within a testing session, they completed two randomly determined conditions (e.g., 2:2, 32:16), each of which lasted for approximately 12 min. Overall, participants completed 360 experimental trials (60 per condition), and the entire experiment lasted for 5 hr. All other details were identical to Experiment 1.

## Results and Discussion

Performance on the task was converted into  $d'$  using the recommendations of Bi (2002), and these data are shown in Figure 6. Four of the participants (DM, RS, TM, and VG) demonstrated one

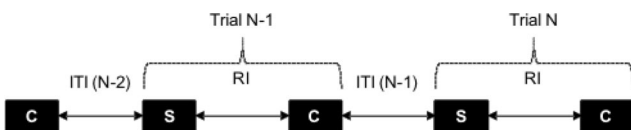


Figure 5. Schematic for trial arrangement in Cowan et al.’s (2001) reassessment of distinctiveness. ITI = intertrial interval; RI = retention interval.

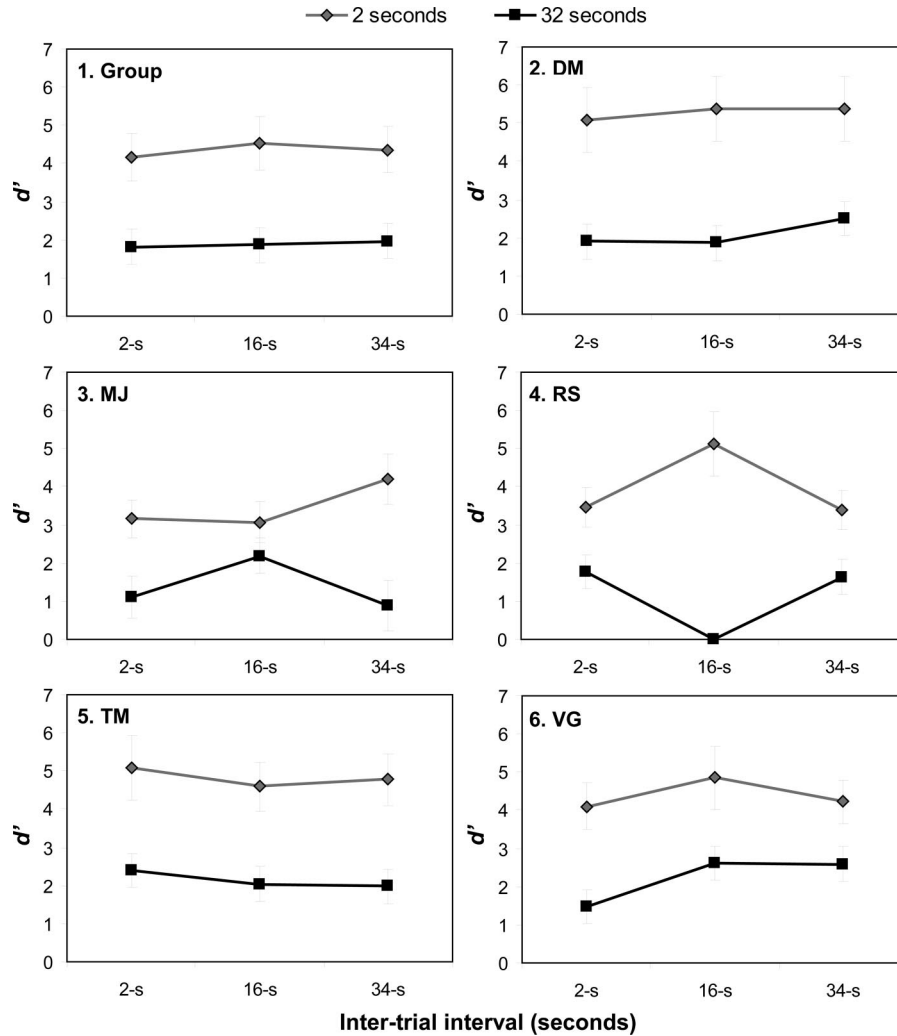


Figure 6. The  $d'$  values in Experiment 2 for each intertrial interval. The gray line shows performance for the 2-s retention interval, and the black line shows performance for the 32-s retention interval. Group data are outlined in Panel 1, and the remaining five panels show  $d'$  for each individual participant. Error bars reflect  $\pm 1$  SE of  $d'$ . No error bar is shown for RS in the 32:16 condition because performance was at chance. Variance of  $d'$  had to be estimated in the analyses for this participant.

or more cases of extreme performance, and therefore all of their scores were converted using the log-linear rule. Still, only 10% of the entire range of  $P_{ss}$  and  $P_{sd}$  values were adjusted in this manner. Figure 6 clearly establishes that discriminatory accuracy was substantially higher when the standard and comparison tones were separated by 2 s rather than by 32 s. This effect persisted regardless of the ITI duration. Indeed, ITI appeared to have relatively little impact upon performance, and temporal isolation rarely seemed to have a beneficial effect. Furthermore, there were cases where temporal isolation resulted in a decrease in task accuracy.

Marascuilo's (1970)  $K$ -signal test was employed to assess these trends. First, the effect of ITI was examined for the 2-s RI (i.e., 2:2 vs. 2:16 vs. 2:34), but no significant results were found for group performance,  $\chi^2(2) = 0.15$ ,  $p = .928$ , or for any of the individual participants: DM,  $\chi^2(2) = 0.08$ ,  $p = .961$ ; MJ,  $\chi^2(2) = 2.02$ ,  $p = .364$ ; RS,  $\chi^2(2) = 3.46$ ,  $p = .177$ ; TM,  $\chi^2(2) = 0.22$ ,  $p = .896$ ; and

VG,  $\chi^2(2) = 0.57$ ,  $p = .752$ . Similarly, there was no effect of ITI within the 32-s RI conditions (32:2 vs. 32:16 vs. 32:34): Group,  $\chi^2(2) = 0.06$ ,  $p = .970$ ; DM,  $\chi^2(2) = 1.28$ ,  $p = .527$ ; MJ,  $\chi^2(2) = 3.53$ ,  $p = .171$ ; RS,  $\chi^2(2) = 1.27$ ,  $p = .530$ ; TM,  $\chi^2(2) = 0.45$ ,  $p = .799$ ; and VG,  $\chi^2(2) = 2.39$ ,  $p = .303$ . Although there was no effect of ITI, the influence of the RI was much more pronounced.  $P_{ss}$  and  $P_{sd}$  values were averaged across ITI to calculate a collapsed  $d'$  for the two RIs (see Figure 7). Sensitivity values in the 2-s and 32-s conditions were then compared with Gourevitch and Galanter's (1967) test. Similar to the time-based forgetting observed in Experiment 1, these analyses revealed significantly higher performance at the shorter gap in all cases: Group,  $Z = 5.42$ ,  $p < .0001$ ; DM,  $Z = 5.05$ ,  $p < .0001$ ; MJ,  $Z = 4.64$ ,  $p < .0001$ ; RS,  $Z = 5.80$ ,  $p < .0001$ ; TM,  $Z = 5.70$ ,  $p < .0001$ ; and VG,  $Z = 4.59$ ,  $p < .0001$ . Further examination revealed that discriminatory accuracy for 2 s was significantly better than 32 s



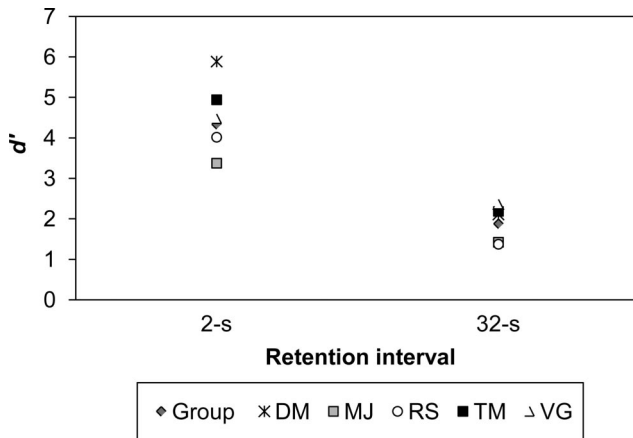


Figure 7. The  $d'$  averaged across intertrial interval for the group and each individual participant at 2- and 32-s retention intervals.

for every corresponding ITI, with just one exception (participant MJ, 2:16 vs. 32:16,  $Z = 1.24$ ,  $p = .215$ ).

These data demonstrate the role of absolute amount of time in determining performance in the *nime* task. They suggest too that the task is uncontaminated by proactive interference or carry-over from prior trials. Each trial appears to “sit alone” in memory space, and therefore what we are observing, when we extend our RI, are processes of forgetting that have previously not been made clear, at least for nonverbal encoding of sound information. We make no claim for other aspects of forgetting of sounds, notably their pitch, and arguably pitch has been the major focus of auditory memory research (Deutsch, 1999). We have suggested previously a role for attentional cueing (T. J. Green & McKeown, 2001; McKeown & Wellsted, 2009), whereby single or pure tones may direct a listener’s attention to their frequency and that this effect may persist across trials. Thus, a listener focusing on the pitch of a pure tone on one trial is likely to still be monitoring that frequency on a subsequent trial (see T. Green & McKeown, 2007, for confirmation of across-trial cueing), increasing greatly the chance of observing trial carry-over in pitch comparisons. However, of course our stimulus comparisons neither encourage, nor benefit from, frequency monitoring. Rather the opposite, as within-stimulus across frequency percepts are demanded in discriminating very fine spectral changes in complex tones (D. M. Green, 1988). Hence, we believe the *nime* offers a clearer window onto auditory memory without carry-over across trials.

However, what about the three-consonant triads of Unsworth et al. (2008)? We imagine, with 21 consonants to play with, these experimental trials shared consonant features on successive trials so that proactive interference should be higher with short ITIs (the authors do not provide sufficient information to evaluate this, unfortunately). That is, we would predict strong trial carry-over effects where successive trials share features through feature overwriting (Mercer & McKeown, 2010b), and this interference should be greater for short delays between trials. There is another difficulty with Unsworth et al.’s study. Participants engaged in a distractor task of repeating aloud a three-digit number during the RI on each trial. To the extent that this activity proactively interferes with performance of verbally encoding a trigram on a sub-

sequent trial, any such effect should be reduced when that trial is followed by a very extended ITI. Therefore, it may not have been reduced temporal distinctiveness that the authors observed following long ITIs but rather reduced carry-over of the distracting activity from the preceding trial following those intervals, and increased feature overwriting following short ITIs (see also, Oberauer & Lange, 2008). Again, we believe that by avoiding the necessity of verbal distraction within the RI, the *nime* is optimum for observing any “real” distinctiveness effects. Of course, we acknowledge there may be studies on distinctiveness in verbal memory using single memoranda that avoid some of the handicaps of Unsworth et al.

## General Discussion

The contribution of the two experiments reported here is to chart the steady forgetting of stimulus memory over a half-minute empty interval, in a situation controlling for proactive interference and active rehearsal. In our introduction, we revisited an old debate on the role of time-based decay in forgetting in immediate memory. The early strategy (Angell & Harwood, 1899) is evident in our approach to settling the debate: We have extended the RI to examine whether the contents of memory appear to lose their fidelity over time. However, research to date has been plagued by two major hurdles. The first is the juggling-like maintenance of activation strength of immediate memory by rehearsal (Nairne, 2002). The second is that one must control for memory-disrupting effects of interfering events before recall. The fact that strategies to deal with each problem run into conflict with one another is acknowledged by theorists (e.g., G. D. A. Brown & Lewandowsky, 2010); that is, to circumvent rehearsal, one must introduce an activity (articulatory suppression) that itself interferes.

A certain amount of ingenuity may be employed in addressing both of these methodological problems. Thus, Lewandowsky et al. (2004) introduced self-paced distractors to block verbal rehearsal into the RI in memory for verbal material (the distractor was the spoken word “super”). Intriguingly, increased numbers of self-paced distractors barely influenced performance even though they increased the RI significantly. Even adding a timed response task to block attentional refreshing of the memoranda (Raye, Johnson, Mitchell, Greene, & Johnson, 2007) during the interval had little effect (Oberauer & Lewandowsky, 2008), although Barrouillet et al. (2011) have questioned the success of this manipulation. Maybe a better approach is to create a situation where active rehearsal or refreshing would be counterproductive. T. Green and McKeown (2007) exploited the fact that detection of a pure tone in noise is enhanced in the presence of a same-frequency pretrial cue tone (T. J. Green & McKeown, 2001). To investigate memory traces in the absence of rehearsal, T. Green and McKeown (2007, Experiment 2) reasoned that, because their listeners were actively attending to a cue tone frequency on trial  $n$ , it was highly unlikely that they were actively rehearsing on trial  $n$  the frequency of the tone on trial  $n - 1$  (and also, presumably such rehearsal would be disruptive to performance). They arranged for the cue tone on trial  $n$  to be at an invalid frequency, and they arranged trials as either trace trials (the signal on trial  $n$  matched the frequency of the cue on trial  $n - 1$ ) or no-trace trials (the signal was at a frequency that had not occurred in at least four previous trials). Signals on trace trials were presented, on average, 5.55 s after the onset of the cue

from the previous trial, and signals on no-trace trials were presented at least 21.55 s after the previous same-frequency cue. Memory without active rehearsal was indicated in that thresholds were lower on recent positive-cue or trace trials than on no-trace trials (the latter did not differ from no-cue control trials). Yet, performance was still poorer on trace trials than trials where the cue on trial  $n$  was itself valid, indicating decay of the memory trace for the previous trial cue. This strategy of examining carry-over of memory traces of stimuli from previous trials, which participants are unlikely to be actively rehearsing, was more recently adopted by Berman et al. (2009) for verbal memory. Instead of recent positive cues, they examined the negative influence of prior trial probe words on response times to current trial target words. They observed persisting negative influences of previous trial probe words on current trial response times. However, unlike the apparent diminishing influence over time of recent positive cues in T. Green and McKeown (2007), suggesting decay, Berman et al. observed *little or no influence* of delay when time from the previous trial was varied from about 1 s to about 10 s.

Importantly, our *nime* task avoids categorical encoding of stimulus information. This is because on “different” trials, the two stimuli differ relatively little in their timbre, determined by relatively small changes to their spectral profile; this spectral manipulation is not enough to cause a category or verbal label change from one stimulus (the “standard” beginning the trial) to the other (the “target” ending the trial). These abstract synthetic sounds also do not lend themselves to verbalization: They cannot be spoken or internally verbally rehearsed. Indeed, we have independent evidence for this, in that reading aloud during the RI barely influences performance on our task (McKeown et al., 2011). The *nime* also dismisses the inconvenience of the establishment of session-long template-like representations by varying stimuli from trial to trial—better to get to the heart of coding for the immediate percept of a stimulus on each trial (i.e., we have avoided category-like “context coding”; cf. Durlach & Braida, 1969). Thus, the frequency range of the two harmonic stimuli on trial  $n$ , determined by their fundamental frequency, is different from that of the two stimuli on trial  $n - 1$ . Importantly, this manipulation of altering the spectral region occupied by our stimuli on successive trials will to a significant degree reduce, or even eliminate, the influence of a form of stimulus-specific proactive influences from prior trials, whereby old trial stimuli overwrite or interact with subsequent stimuli, as demonstrated in our earlier work (Mercer & McKeown, 2010b).

The present data do suggest that memory traces for recent sound stimuli diminish over time, in a situation controlling for verbal rehearsal and proactive interference. This seems to conflict with recent studies of verbal short-term memory that have strongly argued against time-based forgetting (see Lewandowsky et al., 2009), but this may be due to the nature of the stimuli. The majority of recent studies of short-term forgetting have used materials that can be categorized (e.g., letters, digits, or words), but the complex tones employed in the present experiments prevented any categorical encoding. As such, we believe that the memory uncovered in the present studies is a code that conforms to many of the distinguishing characteristics that have been taken to define sensory memory: It appears independent of the direction of attention, it preserves very fine detail of recent stimuli, and it is modality-specific. Yet, it is long-lasting. This last departure from

conventional descriptions of a very brief sensory store lasting a second or so has led us to confront the traditional definition and to suggest (McKeown & Wellsted, 2009; Mercer & McKeown, 2010a) that auditory sensory memory may more properly be viewed as a longer term sensory store, perhaps instantiating the properties of the running comparator revealed through recordings of human evoked potentials in studies of the auditory “mismatch negativity” (see Winkler & Cowan, 2005, for a discussion of the relevance of such evoked potentials studies to settling arguments about the distinctions between sensory and categorical auditory memory). Under this view, the longer sensory auditory store is indeed a context-setter: It is both a detailed record of the most immediate past but forms the immediate context against which upcoming auditory stimulation is recorded.

It therefore appears that there is time-based forgetting of this form of auditory sensory memory, but the exact cause of the information loss we observed remains to be determined. An intriguing possibility is that, like the vanishing impression of a footprint within a dense recovering material such as hard sponge, one form of forgetting within auditory immediate memory is recovery from adaptation in populations of nerve fibers within each stimulated frequency-channel, acting as a TMM of the recent auditory past (McKeown & Wellsted, 2009; Mercer & McKeown, 2010b). Such a view is compatible with an extensive literature on stimulus-specific adaptation revealed in single-cell recordings in the auditory pathways of animals (e.g., Bäuérle, von der Behrens, Kössl, & Gaese, 2011; Ulanovsky, Las, Farkas, & Nelken, 2004) as well as with evidence arising from studies of the auditory mismatch negativity (May & Tiitinen, 2010). However, evidence is also accumulating that synchronization of neural activity within cell assemblies may underlie memory formation and maintenance (see Jutras & Buffalo, 2010, for one recent overview). In particular, low-frequency oscillations in the gamma-band (approximately 30–100 Hz) and theta-band (4–8 Hz)—identified largely through noninvasive electro- or magnetoencephalogram or through intracranial recordings in presurgical patients—are correlated with memory states in humans (e.g., Fell et al., 2001) and animals (e.g., Jutras, Fries, & Buffalo, 2009). The implication is that temporal synchronization of neuronal activity may be uncovering basic mechanisms of the laying down and maintenance of memories for recent stimuli. For auditory stimuli, the relationship between gamma-band oscillations and behavioral performance in memory tasks, involving matching and nonmatching sounds, is established (Kaiser, Heidegger, & Lutzenberger, 2008). More speculative is that synchronous oscillations may underlie binding of stimulus features (Bertrand & Tallon-Baudry, 2000). However, whatever the underlying mechanisms of auditory memory maintenance, studying specifically *nonverbal* memory promises exciting possibilities for settling the long-standing debates about the nature of forgetting.

## References

- Angell, F., & Harwood, H. (1899). Experiments on discrimination of clangs for different intervals of time. *The American Journal of Psychology*, *11*, 67–79. doi:10.2307/1412729
- Barrouillet, P., Portrat, S., Vergauwe, E., Diependaele, K., & Camos, V. (2011). Further evidence for temporal decay in working memory: Reply to Lewandowsky and Oberauer. *Journal of Experimental Psychology*:

- Learning, Memory, and Cognition*, 37, 1302–1317. doi:10.1037/a0022933
- Bäuerle, P., von der Behrens, W., Kössl, M., & Gaese, B. H. (2011). Stimulus-specific adaptation in the gerbil primary auditory thalamus is the result of a fast frequency-specific habituation and is regulated by the corticofugal system. *Journal of Neuroscience*, 31, 9708–9722. doi:10.1523/JNEUROSCI.5814-10.2011
- Berman, M. G., Jonides, J., & Lewis, R. L. (2009). In search of decay in verbal short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 317–333. doi:10.1037/a0014873
- Bertrand, O., & Tallon-Baudry, C. (2000). Oscillatory gamma activity in humans: A possible role for object representation. *International Journal of Psychophysiology*, 38, 211–223. doi:10.1016/S0167-8760(00)00166-5
- Bi, J. (2002). Variance of  $d'$  for the same-different method. *Behavior Research Methods, Instruments & Computers*, 34, 37–45. doi:10.3758/BF03195421
- Bi, J., Ennis, D. M., & O'Mahony, M. (1997). How to estimate and use the variance of  $d'$  from difference tests. *Journal of Sensory Studies*, 12, 87–104. doi:10.1111/j.1745-459X.1997.tb00055.x
- Brown, G. D. A., & Lewandowsky, S. (2010). Forgetting in memory models: Arguments against trace decay and consolidation failure. In S. Della Sala (Ed.), *Forgetting* (pp. 49–75). Hove, England: Psychology Press.
- Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review*, 114, 539–576. doi:10.1037/0033-295X.114.3.539
- Brown, G. S., & White, K. G. (2005). The optimal correction for estimating extreme discriminability. *Behavior Research Methods*, 37, 436–449. doi:10.3758/BF03192712
- Bull, A. R., & Cuddy, L. L. (1972). Recognition memory for pitch of fixed and roving stimulus tones. *Perception & Psychophysics*, 11, 105–109. doi:10.3758/BF03212696
- Clément, S., Demany, L., & Semal, C. (1999). Memory for pitch versus memory for loudness. *The Journal of the Acoustical Society of America*, 106, 2805–2811. doi:10.1121/1.428106
- Cowan, N. (2008). Sensory memory. In H. L. Roediger III (Ed.) & J. Byrne (Vol. Ed.), *Cognitive psychology of memory: Vol. 2. Learning and memory: A comprehensive reference* (pp. 23–32). Oxford, England: Elsevier.
- Cowan, N., Saults, J. S., & Nugent, L. D. (1997). The role of absolute and relative amounts of time in forgetting within immediate memory: The case of tone-pitch comparisons. *Psychonomic Bulletin & Review*, 4, 393–397. doi:10.3758/BF03210799
- Cowan, N., Saults, J. S., & Nugent, L. D. (2001). The ravages of absolute and relative amounts of time on memory. In H. L. Roediger III, J. S. Nairne, I. Neath, & A. Suprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 315–330). Washington, DC: American Psychological Association. doi:10.1037/10394-017
- Crowder, R. G. (1989). Imagery for musical timbre. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 472–478. doi:10.1037/0096-1523.15.3.472
- Demany, L., & Semal, C. (2008). The role of memory in auditory perception. In W. A. Yost, A. N. Popper, & R. F. Fays (Eds.), *Auditory perception of sound sources* (pp. 77–114). New York, NY: Springer-Verlag.
- Demany, L., Trost, W., Serman, M., & Semal, C. (2008). Auditory change detection: Simple sounds are not memorized better than complex sounds. *Psychological Science*, 19, 85–91. doi:10.1111/j.1467-9280.2008.02050.x
- Deutsch, D. (1999). The processing of pitch combinations. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 349–411). New York, NY: Academic Press.
- Dewar, M. T., Cowan, N., & Della Salla, S. (2007). Forgetting due to retroactive interference: A fusion of Müller and Pilzecker's (1900) early insights into everyday forgetting and recent research on anterograde amnesia. *Cortex*, 43, 616–634. doi:10.1016/S0010-9452(08)70492-1
- Durlach, N. I., & Braida, L. D. (1969). Intensity perception: I. Preliminary theory of intensity resolution. *The Journal of the Acoustical Society of America*, 46, 372–383. doi:10.1121/1.1911699
- Durlach, N. I., Braida, L. D., & Ito, Y. (1986). Towards a model for discrimination of broadband signals. *The Journal of the Acoustical Society of America*, 80, 63–72. doi:10.1121/1.394084
- Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review*, 9, 59–79. doi:10.3758/BF03196257
- Fell, J., Klaver, P., Lehnertz, K., Grunwald, T., Schaller, C., Elger, C. E., & Fernandez, G. (2001). Human memory formation is accompanied by rhinal–hippocampal coupling and decoupling. *Nature Neuroscience*, 4, 1259–1264. doi:10.1038/nn759
- Gourevitch, V., & Galanter, E. (1967). A significance test for one parameter isosensitivity functions. *Psychometrika*, 32, 25–33. doi:10.1007/BF02289402
- Green, D. M. (1988). *Profile analysis: Auditory intensity discrimination*. New York, NY: Oxford University Press.
- Green, T. J., & McKeown, J. D. (2001). Capture of attention in selective frequency listening. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 1197–1210. doi:10.1037/0096-1523.27.5.1197
- Green, T., & McKeown, D. (2007). The role of auditory memory traces in attention to frequency. *Perception & Psychophysics*, 69, 942–951. doi:10.3758/BF03193931
- Harris, J. D. (1952). The decline of pitch discrimination with time. *Journal of Experimental Psychology*, 43, 96–99. doi:10.1037/h0057373
- Jutras, M. J., & Buffalo, E. A. (2010). Synchronous neural activity and memory formation. *Current Opinion in Neurobiology*, 20, 150–155. doi:10.1016/j.conb.2010.02.006
- Jutras, M. J., Fries, P., & Buffalo, E. A. (2009). Gamma-band synchronization in the macaque hippocampus and memory formation. *The Journal of Neuroscience*, 29, 12521–12531. doi:10.1523/JNEUROSCI.0640-09.2009
- Kaernbach, C. (2004). Auditory sensory memory and short-term memory. In C. Kaernbach, E. Schröger, & H. Müller (Eds.), *Psychophysics beyond sensation: Laws and invariants of human cognition* (pp. 331–348). Hillsdale, NJ: Erlbaum.
- Kaernbach, C., & Schlemmer, K. (2008). The decay of pitch memory during rehearsal. *The Journal of the Acoustical Society of America*, 123, 1846–1849. doi:10.1121/1.2875365
- Kaiser, J., Heidegger, T., & Lutzenberger, W. (2008). Behavioral relevance of gamma-band activity for short-term memory-based auditory decision-making. *European Journal of Neuroscience*, 27, 3322–3328. doi:10.1111/j.1460-9568.2008.06290.x
- 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. doi:10.1037/0278-7393.21.3.635
- Kinchla, R. A., & Smyzer, F. (1967). A diffusion model of perceptual memory. *Perception & Psychophysics*, 2, 219–229. doi:10.3758/BF03212471
- Lewandowsky, S., Duncan, M., & Brown, G. D. A. (2004). Time does not cause forgetting in short-term serial recall. *Psychonomic Bulletin & Review*, 11, 771–790. doi:10.3758/BF03196705
- Lewandowsky, S., Geiger, S. M., Oberauer, K., & Morrell, D. B. (2010). Turning simple span into complex span: Time for decay or interference from distractors? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36, 958–978. doi:10.1037/a0019764
- Lewandowsky, S., Nimmo, L. M., & Brown, G. D. A. (2008). When temporal isolation benefits memory for serial order. *Journal of Memory and Language*, 58, 415–428. doi:10.1016/j.jml.2006.11.003

- Lewandowsky, S., Oberauer, K., & Brown, G. D. A. (2009). No temporal decay in verbal short-term memory. *Trends in Cognitive Sciences, 13*, 120–126. doi:10.1016/j.tics.2008.12.003
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Cambridge, England: Cambridge University Press.
- Marascuilo, L. A. (1970). Extensions of the significance test for one-parameter signal detection hypotheses. *Psychometrika, 35*, 237–243. doi:10.1007/BF02291265
- Marozeau, J., de Cheveigné, A., McAdams, S., & Winsberg, S. (2003). The dependency of timbre on fundamental frequency. *The Journal of the Acoustical Society of America, 114*, 2946–2957. doi:10.1121/1.1618239
- May, P. J. C., & Tiitinen, H. (2010). Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology, 47*, 66–122. doi:10.1111/j.1469-8986.2009.00856.x
- McKeown, D., Mills, R., & Mercer, T. (2011). Comparisons of complex sounds over extended retention intervals survive reading aloud. *Perception, 40*, 1193–1205. doi:10.1068/p6988
- McKeown, D., & Wellsted, D. (2009). Auditory memory for timbre. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 855–875. doi:10.1037/a0013708
- McLachlan, N., & Wilson, S. (2010). The central role of recognition in auditory perception: A neurobiological model. *Psychological Review, 117*, 175–196. doi:10.1037/a0018063
- Menon, V., Levitin, D. J., Smith, B. K., Lembke, A., Krasnow, B. D., Glazer, D., . . . McAdams, S. (2002). Neural correlates of timbre change in harmonic sounds. *NeuroImage, 17*, 1742–1754. doi:10.1006/nimg.2002.1295
- Mercer, T., & McKeown, D. (2010a). Interference in short-term auditory memory. *The Quarterly Journal of Experimental Psychology, 63*, 1256–1265. doi:10.1080/17470211003802467
- Mercer, T., & McKeown, D. (2010b). Updating and feature overwriting in short-term memory for timbre. *Attention, Perception, & Psychophysics, 72*, 2289–2303.
- Moss, S. M., Myers, J. L., & Filmore, T. (1970). Short-term recognition memory of tones. *Perception & Psychophysics, 7*, 369–373. doi:10.3758/BF03208669
- Nairne, J. S. (2002). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology, 53*, 53–81. doi:10.1146/annurev.psych.53.100901.135131
- Oberauer, K., & Lange, E. B. (2008). Interference in verbal working memory: Distinguishing similarity-based competition, feature overwriting, and feature migration. *Journal of Memory and Language, 58*, 730–745. doi:10.1016/j.jml.2007.09.006
- Oberauer, K., & Lewandowsky, S. (2008). Forgetting in immediate serial recall: Decay, temporal distinctiveness, or interference? *Psychological Review, 115*, 544–576. doi:10.1037/0033-295X.115.3.544
- Raye, C. L., Johnson, M. K., Mitchell, K. J., Greene, E. J., & Johnson, M. R. (2007). Refreshing: A minimal executive function. *Cortex, 43*, 135–145. doi:10.1016/S0010-9452(08)70451-9
- Ricker, T. J., & Cowan, N. (2010). Loss of visual working memory within seconds: The combined use of refreshable and non-refreshable features. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 1355–1368. doi:10.1037/a0020356
- Ricker, T. J., Cowan, N., & Morey, C. C. (2010). Visual working memory is disrupted by covert verbal retrieval. *Psychonomic Bulletin & Review, 17*, 516–521. doi:10.3758/PBR.17.4.516
- Sams, M., Hari, R., Rif, J., & Knuutila, J. (1993). The human auditory sensory memory trace persists about 10 sec: Neuromagnetic evidence. *Journal of Cognitive Neuroscience, 5*, 363–370. doi:10.1162/jocn.1993.5.3.363
- Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117*, 34–50. doi:10.1037/0096-3445.117.1.34
- Ulanovsky, N., Las, L., Farkas, D., & Nelken, I. (2004). Multiple time scales of adaptation in auditory cortex neurons. *The Journal of Neuroscience, 24*, 10440–10453. doi:10.1523/JNEUROSCI.1905-04.2004
- Unsworth, N., Heitz, R. P., & Parks, N. A. (2008). The importance of temporal distinctiveness for forgetting over the short term. *Psychological Science, 19*, 1078–1081. doi:10.1111/j.1467-9280.2008.02203.x
- Visscher, K. M., Kahana, M., & Sekuler, R. (2009). Trial-to-trial carryover in auditory short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 46–56. doi:10.1037/a0013412
- Whipple, G. M. (1901). An analytic study of the memory image and the process of judgment in the discrimination of clangs and tones. *The American Journal of Psychology, 12*, 409–457. doi:10.2307/1412145
- Winkler, I., & Cowan, N. (2005). From sensory to long-term memory: Evidence from auditory memory reactivation studies. *Experimental Psychology, 52*, 3–20. doi:10.1027/1618-3169.52.1.3

Received August 6, 2011

Revision received December 12, 2011

Accepted December 19, 2011 ■