

## Learning-related shifts in generalization gradients for complex sounds

MATTHEW G. WISNIEWSKI, BARBARA A. CHURCH, AND EDUARDO MERCADO III  
University at Buffalo, State University of New York, Buffalo, New York

Learning to discriminate stimuli can alter how one distinguishes related stimuli. For instance, training an individual to differentiate between two stimuli along a single dimension can alter how that individual generalizes learned responses. In this study, we examined the persistence of shifts in generalization gradients after training with sounds. University students were trained to differentiate two sounds that varied along a complex acoustic dimension. The students were subsequently tested on their ability to recognize a sound that they had experienced during training when it was presented among several novel sounds varying along this same dimension. Peak shift was observed in Experiment 1, in which generalization tests immediately followed training, and in Experiment 2, in which the tests were delayed by 24 h. These findings further support the universality of generalization processes across species, modalities, and levels of stimulus complexity. They also raise new questions about the mechanisms underlying learning-related shifts in generalization gradients. The sound stimuli from this study are available as .wav files from <http://lb.psychonomic-journals.org/content/supplemental>.

*Generalization* is the tendency for organisms to differentially judge novel stimuli as predictive of a learned consequence on the basis of the novel stimuli's similarity to stimuli experienced during learning (Shepard, 1987). The phenomenon is often studied experimentally by training a subject to respond to one stimulus (S+). Generalization is then measured by removing feedback, presenting novel stimuli, and measuring responses to stimuli that are similar to the S+. In a classic example, pigeons trained to respond to a key illuminated with a 580-nm light also responded to other wavelengths when reinforcement for their responses was absent; the number of responses to newly presented wavelengths lessened with decreasing similarity to the conditioned stimulus (Guttman & Kalish, 1956). This phenomenon is seen in animals ranging from invertebrates (Cheng, 1999, 2000, 2002) to humans (reviewed by Thomas, 1993) with various testing strategies and dimensions (Shepard, 1987). Because of its ubiquity, generalization was described by Pavlov (1927) as being a fundamental associative process and by Shepard as psychology's first law. The universality of generalization likely reflects the fact that an organism's ability to react appropriately to new environmental conditions often depends on its capacity to predict possible outcomes on the basis of prior experience (Ghirlanda & Enquist, 2003; Shepard, 1987).

Despite the pervasiveness and importance of generalization across species, experimental studies of this phenomenon in humans and those in nonhuman animals have proceeded somewhat independently. Processes postulated to underlie generalization by humans, such as identifying relational cues or rules (Ahn & Medin, 1992; Doll &

Thomas, 1967; Helson, 1964; Imai & Garner, 1965; Shepard, Hovland, & Jenkins, 1961; Thomas, 1993; Thomas & Bistey, 1964; Verbeek, Spetch, Cheng, & Clifford, 2006; Wattenmaker, 1992), using abstract strategies (Gunzelmann, 2008; Rodrigues & Murre, 2007), and language (Fagot, Goldstein, Davidoff, & Pickering, 2006; Purtle, 1973; Shepard et al., 1961), are often difficult or impossible to investigate across species. Studies of nonhuman animals have instead focused on measuring the form of generalization gradients, context effects, dimensional variety, and training procedures (for reviews, see Ghirlanda & Enquist, 2003; Honig & Urcuioli, 1981; Purtle, 1973). There are some comparable data sets from human participants (Cheng, Spetch, & Johnston, 1997; Doll & Thomas, 1967; Spetch & Cheng, 1998; Thomas & Bistey, 1964; Thomas & Mitchell, 1962), but they are relatively rare. Although differences in the processes underlying generalization gradients in humans and nonhumans undoubtedly exist, many fundamental mechanisms are likely shared (Cheng, 2002; Ghirlanda & Enquist, 1999, 2003; Shepard, 1987).

A basic process in all animals that strongly impacts generalization is learning. For example, Jenkins and Harrison (1960) found that pigeons reinforced after responses to a tone displayed narrower generalization gradients as they experienced more training blocks. Another effect of learning on generalization, described as the *peak shift effect* (Hanson, 1959; Purtle, 1973; Spence, 1937), typically occurs after intradimensional discrimination training in which one stimulus is reinforced (S+), and another stimulus is not (S-). When generalization tests are given after such training, subjects mostly respond not to S+, as one might expect, but to one of the unconditioned test stimuli.

These shifts are often accompanied by a similar shift in gradient mean and the overall proportion of responding (Bizo & McMahan, 2007; Newlin, Rodgers, & Thomas, 1979; Thomas & Bistey, 1964; Thomas, Mood, Morrison, & Wiertelak, 1991). In a classic example, pigeons trained on a wavelength discrimination between 550-nm (S+) and 560-nm (S-) light responded most to shorter wavelengths, such as 540 nm, in a generalization test (Hanson, 1959). Such learning-related gradient shifts have been observed in bees (Lynn, Cnaani, & Papaj, 2005), horses (Dougherty & Lewis, 1991), rats, goldfish, guinea pigs, chickens, pigeons, and humans (for a review, see Purtle, 1973), and with various methods such as category learning (McLaren, Bennett, Guttman-Nahir, Kim, & Mackintosh, 1995; McLaren & Mackintosh, 2002; Wills & Mackintosh, 1998), varying distributions of test stimuli (Bizo & McMahan, 2007; Helson & Avant, 1967; Spetch, Cheng, & Clifford, 2004; Thomas & Bistey, 1964), and training with multiple S+ stimuli (Galizio & Baron, 1979; White & Thomas, 1979). The effect occurs for stimuli varying in brightness (Newlin et al., 1979; Thomas et al., 1991; Thomas, Ost, & Thomas, 1960; White & Thomas, 1979), facial characteristics (Lewis & Johnston, 1999; McLaren & Mackintosh, 2002; Spetch et al., 2004), spatial location (Cheng & Spetch, 2002), pitch (Galizio, 1985; Galizio & Baron, 1979), line tilt (Spetch et al., 2004; Winton & Beale, 1971), floor tilt (Ricchio, Urda, & Thomas, 1966; Thomas & Burr, 1969; Thomas & Lyons, 1968), numerosity (Honig & Stewart, 1993), and motor movement (Dickinson & Hedges, 1986). As with other generalization phenomena, the processes postulated to give rise to the peak shift effect have been different for humans and nonhumans, despite the effect's apparent universality.

In nonhumans, Spence's (1937) theory of gradient summation has been the prevailing explanation. Spence proposed that after training, an excitatory gradient surrounds the S+ stimulus, and an inhibitory gradient surrounds the S-. The peak shift is proposed to result from the summation of these gradients. Specifically, although the S+ is located at the peak of the excitatory gradient, this peak overlaps with part of the inhibitory gradient. As a result, a location displaced from S+ that still has part of the excitatory gradient, but less of the inhibitory gradient, ends up with a higher total excitatory value. For humans, a more commonly cited explanation is the adaptation level account. According to adaptation level theory, people develop a psychological average of the presented stimuli during discrimination training, called an *adaptation level* (Thomas, 1993). Incoming exemplars are then judged on the basis of this learned prototype, and the participants respond to the prototype plus  $x$  units during generalization tests. When the range and distribution of stimuli change from training to generalization tests, the adaptation level also changes, but the rule for responding remains the same. Consequently, use of the rule learned during training produces peak responding to a novel stimulus. Both of these possibilities fail to explain some instances of peak shift (Bizo & McMahan, 2007; Cheng & Spetch, 2002; Doll & Thomas, 1967; Hanson, 1959; Lazareva, Wasserman, & Young, 2005; McLaren et al., 1995;

McLaren & Mackintosh, 2002; Thomas, 1993; Thomas & Bistey, 1964; Thomas et al., 1991), but they have been able to explain other aspects of generalization (Helson, 1964; Spence, 1937; Thomas, 1993). For this reason, it is worthwhile to study peak shift in ways that control for the effects of inhibitory gradients and the stimulus range. By taking this approach, one limits the possibility that aspects of Spence's theory or adaptation level (Helson, 1964; Thomas, 1993) may mask alternative mechanisms underlying shift that are common across species.

One other factor that may have contributed to disparate theoretical accounts of learning-related shifts is an emphasis on species differences. Traditionally, studies of the peak shift effect have involved simple stimuli varying along a single dimension. Such simple stimulus sets may be generalized differently by humans and other animals, because humans can develop verbal response rules for differentiating such stimuli (Ahn & Medin, 1992; Imai & Garner, 1965; Shepard et al., 1961; Wattenmaker, 1992). More recently, experiments in which participants were trained to discriminate between two faces that morphed along a complex visuospatial dimension (Lewis & Johnston, 1999; McLaren & Mackintosh, 2002; Spetch et al., 2004) established that peak shift also occurs with multidimensional stimuli. As a result, questions have been raised about whether explanations necessitating response rules can explain shift with stimuli of greater complexity (Spetch et al., 2004). Further investigation into peak shift with other complex stimuli is needed before conclusions can be reached.

In the present experiments, we further explored the generality of learning-related shifts in generalization gradients by measuring responses in individuals trained to identify temporally dynamic sounds varying along a complex acoustic dimension. Specifically, participants learned to identify sounds in which frequency was modulated periodically at a fixed rate. There appear to be no prior reports of a peak shift effect in humans after training with temporally dynamic stimuli in any modality or after training with multidimensional acoustic stimuli. The first aim of these experiments was thus to assess whether learning-related shifts in generalization gradients similar to those observed after training with complex visual stimuli (e.g., faces) are also observed after training with complex acoustic stimuli. In addition, we controlled for changes in adaptation level between training and testing, as well as for asymmetries in excitatory or inhibitory gradients resulting from differential reinforcement. Observations of learning-related shifts with these controls in place may provide new clues about the mechanisms underlying generalization.

We also explored the time course of the peak shift phenomenon. Aside from a few studies with pigeons (Moye & Thomas, 1982; Spence, 1937; Thomas et al., 1960; Thomas et al., 1985), the time course of peak shift retention has not been explored. Tests of forgetting suggest that as time passes after training, novel stimuli become more likely to elicit the same responses as trained stimuli (reviewed by Bouton, Nelson, & Rosas, 1999; Ricchio, Ackil, & Burch-Vernon, 1992). Similarly, generalization

gradients tend to flatten over time in both humans and nonhumans. These results predict that shifts in generalization might also vary over time. However, Spence (1937) and Moye and Thomas (1982) reported peak shift effects after a 24-h delay between training and test, and Thomas et al. (1960) showed that this effect persisted for at least 3 weeks. Is peak shift similarly stable in humans? We addressed this question by inserting a 24-h delay between training and testing.

Knowing how long learning-related shifts in gradients last can provide insights into their underlying mechanisms (Blough, 1975). If peak shift effects do not persist in humans, short-term mechanisms such as attentional shifts may be sufficient to explain the effect. This finding would support the current assumption of dissociations between generalization in humans and other animals. If the effect is still present after a 24-h delay, however, accounts based on temporary adjustments of stimulus processing or decision criteria are untenable, and additional similarities in mechanisms across species would be revealed.

## EXPERIMENT 1

Experiment 1 was conducted to determine whether peak shift would be observed in participants trained to distinguish complex sounds. We held objective adaptation levels constant between training and testing and controlled for differences between excitatory and inhibitory gradients by using a two-choice task. For ease of discussion and to remain consistent with the literature, we refer to the stimuli experienced during training as S+ and S−, even though the participants received equal reinforcement for responses to both stimuli. The participants in the identification training condition were instructed to respond to S+ on every trial during training. The participants in

the discrimination training condition responded differentially to S+ and S− during training. Both groups were then tested with sounds similar to S+ and S− and were asked to identify the S+. We hypothesized that after identification training, generalization gradients would peak at S+ but that after discrimination training, gradients would peak at a stimulus other than S+. Specifically, we predicted that the generalization gradients of the discrimination training group would display a peak proportion of S+ responses at a location displaced from S+ in a direction further from S−.

## Method

**Participants.** Forty-four introductory psychology students at the University at Buffalo, SUNY, participated in the study for partial course credit. Three participants were dropped from the identification training group. One was dropped because of a computer error and the other 2 for clearly ignoring instructions to make only 15% S+ responses during the generalization phase. The criterion for elimination was an S+ response proportion in excess of 50%. Five participants were dropped from the discrimination training group, because they failed to reach the training criterion of 70% correct. The 70% correct criterion guaranteed that we were observing people who learned the discrimination and who were paying attention to the task. We felt that a difficult discrimination was desirable to ensure that learning was required. A total of 36 participants, 18 in each training condition, were used for the data analyses.

**Stimuli and Apparatus.** The sounds used in this study were 1-sec trains of frequency-modulated (FM) tonal sweeps that varied in repetition rate (see Figure 1)<sup>1</sup>; all sweep trains were generated using MATLAB 6.5. Individual sweeps increased in frequency from 500 to 4000 Hz. Thus, all stimuli spanned the same broad range of frequencies and were of the same duration. There were several advantages to using FM sweep trains. First, the participants were unlikely to have heard the stimuli prior to the experiment. Second, sweep trains generate predictable cortical responses in rats (Orduña, Mercado, Gluck, & Merzenich, 2005) and owl monkeys (deCharms, Blake, & Merzenich, 1998), suggesting that they likely do so in other

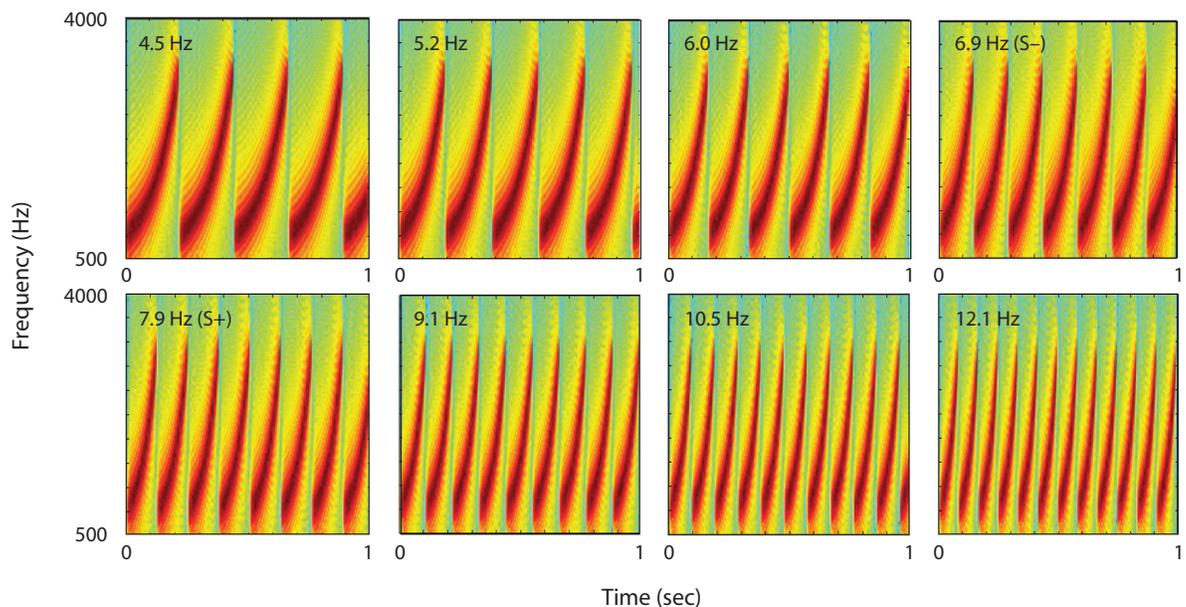


Figure 1. Spectrograms of the stimuli used in Experiments 1 and 2.

mammals, including humans. Finally, the structure of these sounds allows them to be varied independently along several dimensions, including repetition rate, modulation rate, modulation direction, and range of modulation (Mercado, Orduña, & Nowak, 2005).

The repetition rates used were: 4.0, 7.9, and 14 Hz, used for pre-training; 6.9 and 7.9 Hz, used in training; and 4.5, 5.2, 6.0, 6.9, 7.9, 9.1, 10.5, and 12.1 Hz, used for testing. The rates other than those used in pretraining were selected so that the repetition rate of each sound was approximately 15% greater than the prior sound. The increase in rate not only increased the number of sweep repetitions per second, but also altered the length and steepness of each sweep. Therefore, these sounds differed on a number of acoustic dimensions.

The sounds were presented and the responses collected using DMDX experimental software (Forster & Forster, 2003) running on HP Pavilion a300n IBM-compatible desktop computers. The participants heard the sounds at a normal conversational volume through Audio-Technica ATH-M40fs headphones and indicated their responses using the shift keys on a keyboard.

**Design and Procedure.** In Experiment 1, we employed a  $2 \times 8$  mixed factorial design. The between-participants factor (training condition) had two levels (identification and discrimination), and the within-participants factor (repetition rate) had eight levels. One of the dependent measures was the proportion of times that a participant produced a response indicating that S+ had occurred after hearing a particular sound. A second dependent measure was the mean of the participants' obtained gradients.

All of the participants were instructed to respond by pressing the right shift key on a computer keyboard, labeled "8," if they heard one sound (with a repetition rate of 7.9 Hz) and to respond with the left shift key, labeled "NOT 8," if they heard a slower or faster sound. The participants were told to guess if they were unsure of whether the sound that they had heard was at a rate of 8. Each participant engaged in a pretraining task, training, a pretest reminder, and a testing phase. The pretraining task consisted of six trials meant to assist the participants in understanding what they were to do in the task. During these six trials, the participants were presented with four sounds not heard in the training or test (two 4 Hz and two 14 Hz) and two 7.9-Hz S+ sounds. In each pretraining trial, the participants were given the correct answer before making their response ("NOT 8" or "8").

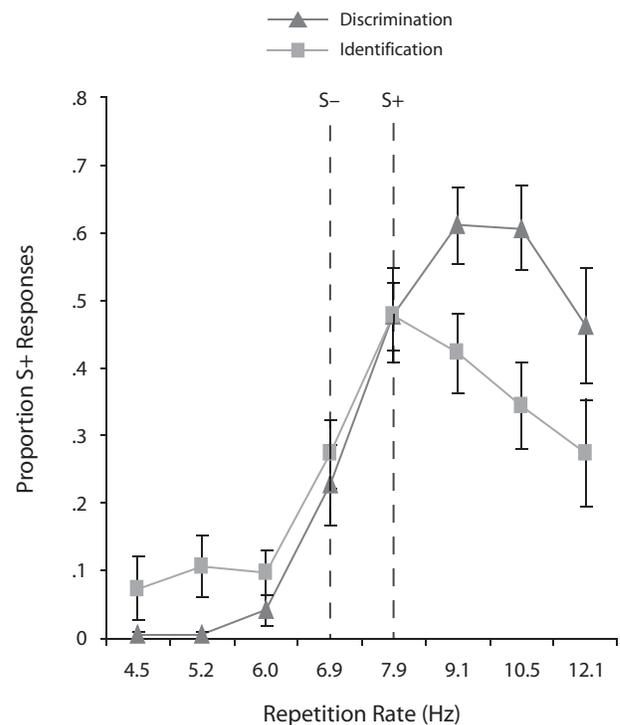
During the training phase, the identification group received 20 presentations of the (S+) stimulus, and the discrimination group received 20 S+ stimuli and 20 S- stimuli. The training stimuli for the discrimination group were presented in pseudorandom order, so that no more than 5 of the same sounds appeared consecutively. The identification group heard 20 consecutive presentations of S+. The participants were prompted to make a response during each trial and were given feedback after their responses. The word "correct" appeared on the screen after a correct response, along with the recorded reaction time in milliseconds. If an incorrect response was made, the word "wrong" appeared on the screen. After training was completed, the participants were instructed that during the next phase of the experiment, only 15% of the sounds would require an "8" response, and the remaining 85% of sweeps would require a "NOT 8" response. The participants were also reminded that their task was to respond "NOT 8" to any sounds that were slower or faster than a rate of 8. These instructions constituted the pretest reminder. The participants then hit the space bar to start the test. During the test, the testing values (4.5, 5.2, 6.0, 6.9, 7.9, 9.1, 10.5, and 12.1 Hz) were presented in pseudorandom order, so that the same stimulus did not occur more than twice in a row. The participants were not given feedback for their responses. Each stimulus was presented 12 times, for a total of 96 test trials.

## Results

The results from averaging the proportions of responding across participants for the discrimination and identification groups are shown in Figure 2.

All statistical tests were two tailed and employed an alpha level of .05. A  $2 \times 8$  ANOVA was performed, with training condition as the between- and repetition rate as the within-participants factor. The analysis revealed a significant main effect of repetition rate [ $F(7,238) = 29.766, p < .001, \eta_p^2 = .467$ ], which shows that some sounds elicited more S+ responding than others. The main effect of condition was not significant, indicating that a difference in overall amount of S+ responding was not detected. There was also a significant repetition rate  $\times$  condition interaction [ $F(7,238) = 3.674, p = .001, \eta_p^2 = .098$ ]. This interaction indicates that the discrimination and identification training groups differed in regards to the proportion of S+ responses that they made to different repetition rate values.

Mean response gradients were calculated and compared in order to determine whether the interaction reflected a peak shift in the discrimination group. To calculate the mean gradients, the repetition rate values were rank-ordered as follows: (1) 4.5 Hz, (2) 5.2 Hz, (3) 6 Hz, (4) 6.9 Hz, (5) 7.9 Hz, (6) 9.1 Hz, (7) 10.5 Hz, and (8) 12.1 Hz. The means of the responding gradients were obtained by multiplying S+ responses by their rank. This value was then divided by the total number of S+ responses. The 7.9-Hz S+ sweep had a rank of 5. The mean gradients were  $M = 6.09, SD = 0.49$ , for the discrimination group;  $M = 5.36, SD = 1.01$ , for the identification group. A planned comparison independent  $t$  test revealed



**Figure 2.** The average proportion of S+ responses to each repetition rate value for the discrimination and identification groups in Experiment 1. Error bars represent one standard error of the mean.

that the means of these groups differed significantly [ $t(34) = 2.69, p = .011$ , Cohen's  $d = 0.93$ ]. To determine whether either of the means was significantly different from the S+ value, we conducted one-sample  $t$  tests on each condition's mean gradient with 5 (S+) as the comparison value. The discrimination group's mean of 6.09 was significantly different from 5, revealing a peak shift [ $t(17) = 9.25, p < .001$ , Cohen's  $d = 2.39$ ]. The control group's mean of 5.36 was not significantly different from 5 [ $t(17) = 1.53, p = .143$ , Cohen's  $d = 0.37$ ].<sup>2</sup>

### Discussion

The results of Experiment 1 display a classic peak shift, even though the factors important for either adaptation level or Spence's (1937) theory to predict peak shift were eliminated. Explanations based on changes in adaptation level cannot account for the observed shift, because the relational responding learned in training, if used later in the testing phase, should have yielded peak responding at the S+. In this case, the participants would have learned to respond to some average of the stimuli, presumably between 6.9 and 7.9 Hz, plus some value. Because this average was objectively the same in both training and testing, adding the same value throughout the session would have yielded responding at a consistent location. This result is consistent with those of Spetch et al. (2004), in which the range effects with multidimensional visual stimuli was controlled for. No evidence that adaptation level impacted learning-related shifts was seen (see also Cheng & Spetch, 2002). Spence's theory is also unable to explain this shift, because the participants in the discrimination training group were reinforced equally for responses to both discriminative stimuli. Neither stimulus should have been associated with inhibition, because correct responses in the presence of either stimuli were followed with a "Correct" feedback message. The inhibition necessary for Spence's theory to predict shift was thus not present.

It is also clear that the results do not reflect the participants' counting the number of repetitions. At lower repetition rate values, these stimuli are countable, but with the repetition rate values at or above the S+/S- counting is extremely difficult, if not impossible. There were a number of people who failed to meet criterion in training. If counting had been employed, we would not have seen this attrition. In addition, if the participants could accurately count the majority of the repetition values, there would be no expected difference between the identification and discrimination participants and no peak shift.

Alternative theoretical frameworks that emphasize the representational structure of stimuli provide a more plausible account of the present results. These frameworks assume that a stimulus is represented as graded activation over a set of elements (Blough, 1975; Ghirlanda & Enquist, 1999; McLaren & Mackintosh, 2002; Thorndike, 1932). Similar stimuli, such as those that vary along a particular dimension, share elements; the more similar two stimuli are, the more elements they share. Elemental explanations account for learning-related gradient shifts by assuming

that certain elements of a stimulus that are predictive of the correct response become more salient than elements that are not predictive. This account has previously been used to explain generalization after training with static visual images. For the elemental explanation to apply to the present data, one must assume that elements of stimulus representations include dynamic features and that variations over time activate overlapping sets of elements in the same way as different static features do.

To our knowledge, Experiment 1 is the first to show peak shift in humans after training with stimuli that vary over time. This finding is consistent with recent reports of peak shift effects in humans trained to distinguish faces varying along a continuum (Lewis & Johnston, 1999; McLaren & Mackintosh, 2002; Spetch et al., 2004) and in pigeons trained to discriminate multi-item visual displays (Honig & Stewart, 1993; Wills & Mackintosh, 1998). The present data afford more direct comparisons across species than do studies of humans generalizing across faces, however, because faces are especially salient stimuli for humans (Jitsumori & Makino, 2004), and the participants had an enormous amount of prior experience distinguishing faces prior to training. In contrast, the participants were unlikely to have experienced FM sweep trains before participating in this experiment, and the sounds had no natural significance. Observations of a learning-related shift after training with these sounds thus increase confidence in the global nature of this phenomenon.

### EXPERIMENT 2

In Experiment 2, we assessed whether the peak shift effect observed in Experiment 1 would persist when there was a significant delay interval between discrimination training and generalization testing. Specifically, we planned to replicate Experiment 1 in a zero-delay condition and then contrast these results with those of participants who experienced a 24-h delay between training and testing. Thomas and colleagues (Moye & Thomas, 1982; Thomas & Burr, 1969; Thomas et al., 1960; Thomas et al., 1985) reported that pigeons showed the peak shift effect after this and longer delays. Comparable tests have not been performed in humans. Identifying similarities and differences in generalization across species can clarify whether the proposed dichotomy between mechanisms of human and nonhuman peak shift is justified. In addition, information regarding the nature and function of learning-related gradient shifts may be obtained by observing its time course. For example, if the peak shift effect disappears after a delay, short-term attentional shifts might be a plausible explanation for the effect. If peak shift remains after the 24-h delay, a long-term learning process that is common across species may be more probable. Given Thomas and colleagues' (Moye & Thomas, 1982; Thomas & Burr, 1969; Thomas et al., 1960; Thomas et al., 1985) findings and the results of Experiment 1, we hypothesized that both (with and without delay) discrimination groups would show a classic peak shift and that the identification groups would not.

## Method

**Participants.** Sixty-eight introductory psychology students at the University at Buffalo, SUNY, participated in the study to partially fulfill their course requirements. Seven participants from the discrimination groups tested without delay and 3 from groups tested after a delay were dropped because of their failure to reach the training criterion of 70% correct. To create equal sample sizes, 3 participants from both the identification without delay and the identification with delay groups were dropped randomly and blindly. The total number of participants in each group was 13 after these exclusions.

**Stimuli and Apparatus.** The stimuli and apparatus were the same as those used in Experiment 1.

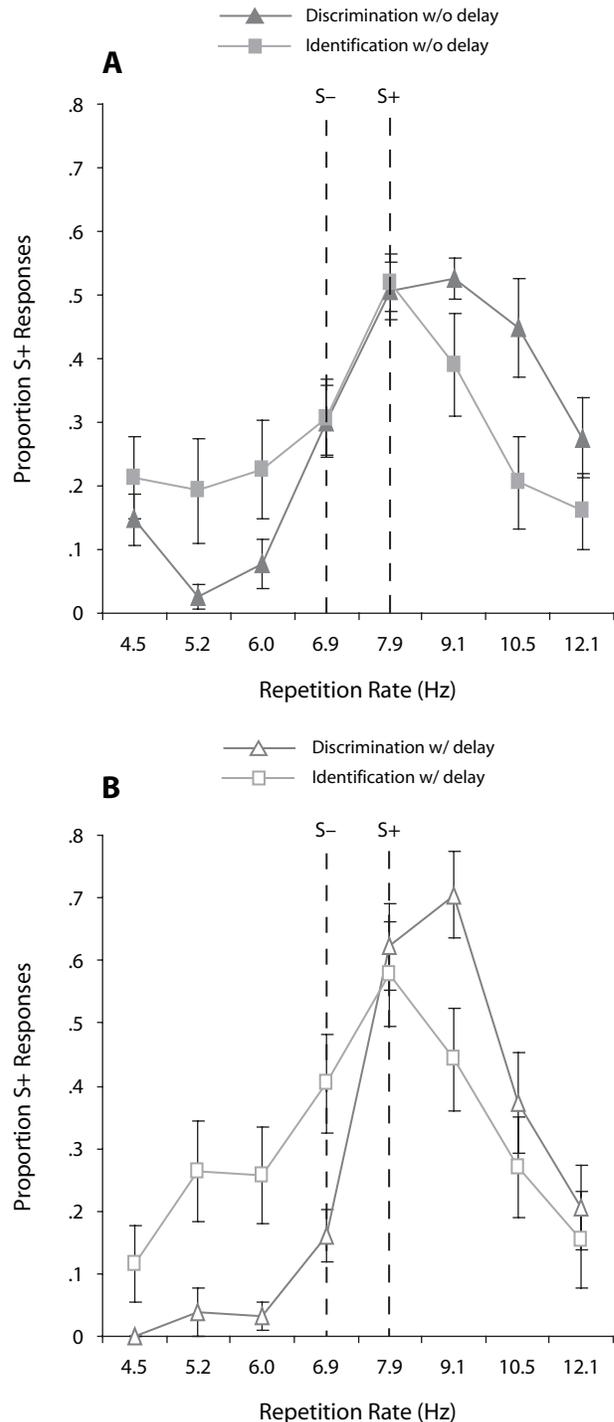
**Design and Procedure.** A  $2 \times 2 \times 8$  mixed factorial design was used. The between-participants factor of training condition and the within-participants factor of repetition rate were the same as in Experiment 1. The second between-participants factor (delay time) had two levels (0 and 24 h). The 0-h delay was an exact replication of the delay in Experiment 1. The dependent measures were the same as those in Experiment 1.

The procedure was also the same as those in Experiment 1, except that instead of proceeding to the generalization period after training, the discrimination and identification training groups with the 24-h delay were told to come back the next day at exactly the same time. Upon returning to the lab, 24 h later, these participants were given the pretest reminder and then the generalization test.

## Results

As in Experiment 1, the dependent measures were the proportion of S+ responses to each particular repetition rate value and the mean gradient of responding. The mean of the participants' responses for each group at each repetition rate value is shown in Figure 3.

All statistical tests were two tailed and used an alpha level of .05. A  $2 \times 2 \times 8$  ANOVA was conducted. The between-participants factors were training condition and delay condition. The within-participants factor was repetition rate. A significant main effect of repetition rate, demonstrating a difference in responding elicited by the repetition rate values, was found [ $F(7,336) = 28.860, p < .001, \eta_p^2 = .375$ ]. The ANOVA also revealed a significant repetition rate  $\times$  condition interaction [ $F(7,336) = 5.978, p < .001, \eta_p^2 = .111$ ]. All other main effects and interactions were not significant ( $F_s < 2$ ). We performed two separate  $2 \times 8$  ANOVAs, with condition as the between- and repetition rate as the within-participants factor to be cautious in accepting peak shift effects at both levels of delay. The ANOVA conducted with the discrimination without delay and the identification without delay groups yielded a significant main effect of repetition rate [ $F(7,168) = 28.785, p < .001, \eta_p^2 = .365$ ] and a significant condition  $\times$  repetition rate interaction [ $F(7,168) = 5.962, p < .001, \eta_p^2 = .107$ ]. The main effect of condition was not significant ( $F < 2$ ), showing the same overall level of responding. The main effect of repetition rate illustrates that different repetition rate values elicited different proportions of S+ responding. The significant repetition rate  $\times$  condition interaction supports a difference between the S+ responding of the training conditions elicited by the presentation of each sound. The ANOVA performed on the discrimination with delay and identification with delay groups also revealed a significant main effect of repetition rate [ $F(7,168) = 17.287, p < .001, \eta_p^2 = .419$ ] and a significant condition  $\times$  repetition rate



**Figure 3.** The average S+ response proportions to each repetition rate value for the discrimination and identification groups in Experiment 2 (A) without a delay between training and test and (B) with a 24-h delay between training and testing. Error bars represent one standard error of the mean.

interaction [ $F(7,168) = 3.691, p = .001, \eta_p^2 = .133$ ], illustrating the same differences as the previous subcomparison ANOVA. The main effect of condition was not significant ( $F < 2$ ).

Again, comparisons of the mean gradients were used to determine whether peak shifts in the discrimination conditions caused the significant interactions. The stimuli were rank ordered, as was done in Experiment 1. The mean gradients of the groups were  $M = 5.43$ ,  $SD = 0.55$ , for the discrimination without delay group;  $M = 4.7$ ,  $SD = 1.14$ , for the identification without delay;  $M = 5.7$ ,  $SD = 0.81$ , for discrimination with delay; and  $M = 4.7$ ,  $SD = 1.12$ , for identification with delay. A 2 (training condition)  $\times$  2 (delay) ANOVA for the means revealed a significant main effect of condition [ $F(1,48) = 10.59$ ,  $p = .002$ ,  $\eta_p^2 = .181$ ]. The main effect of delay and the delay  $\times$  condition interaction were not significant ( $F < 2$ ).

Planned comparison  $t$  tests revealed significant differences between the discrimination without delay and identification without delay groups [ $t(24) = 2.08$ ,  $p = .048$ , Cohen's  $d = 0.85$ ] and between discrimination with delay and identification with delay [ $t(24) = 2.50$ ,  $p = .019$ , Cohen's  $d = 1.02$ ]. One-sample  $t$  tests revealed that the means for the discrimination without delay [ $t(12) = 2.80$ ,  $p = .016$ , Cohen's  $d = 0.85$ ] and the discrimination with delay [ $t(12) = 3.027$ ,  $p = .011$ , Cohen's  $d = 0.94$ ] groups were significantly above 5 (S+), indicating a peak shift. The  $t$  tests performed on the identification without delay and identification with delay groups did not reach significance ( $t < 1$ ), showing that the identification groups' means were not significantly displaced from 5 (S+).<sup>3</sup>

## Discussion

The findings of Experiment 2 replicated those of Experiment 1 and also showed that the peak shift effect remained when there was a 24-h delay between discrimination training and generalization testing. Given that training non-human animals involves significantly more exposures to a stimulus dimension and the explanatory dichotomy in the literature, it was possible that peak shift with humans might not persist as was previously reported in pigeons. On the contrary, we found that training humans for only 40 trials was sufficient to induce a long-lasting effect. This finding suggests the possibility that similar learning mechanisms may underlie peak shift in humans and other animals. In addition, although prior generalization research in both humans and nonhumans has suggested that gradients flatten over time, and several theories of generalization assume this gradient change (Estes, 1997; Nairne, 1991; Riccio et al., 1992), our data give no indications of delay-dependent changes in generalization for either training condition.

## GENERAL DISCUSSION

In this study, participants learned to press particular keys after hearing dynamic acoustic stimuli. Having learned this task, they generalized learned responses to novel stimuli varying along a complex acoustic dimension. The individuals that learned to respond to a single sound responded most to the sound experienced during training and less to novel sounds. In contrast, the individuals trained to respond to two sounds and then tested on their ability to recognize one of the sounds showed a generalization gradient that was shifted. Peak respond-

ing was to a novel stimulus rather than to either of the sounds experienced during training. This peak shift was evident immediately after training (Experiment 1), as well as 24 h after training (Experiment 2). These findings are consistent with those of numerous past reports of a peak shift effect after discrimination learning, but extend this phenomenon to temporally dynamic stimuli. This is also, to our knowledge, the first demonstration in humans that learning-related shifts in generalization gradients can persist for at least a day.

Recent studies have convincingly demonstrated peak shift effects after discrimination training with complex static visual stimuli, including sets of icons (Honig & Stewart, 1993; Wills & Mackintosh, 1998), spatial layouts (Cheng & Spetch, 2002), faces (Lewis & Johnston, 1999; McLaren & Mackintosh, 2002; Spetch et al., 2004), and various other naturalistic images (Ghirlanda & Enquist, 2003; Lynn et al., 2005). Collectively, these studies strongly suggest a common mode of visual discrimination learning that impacts stimulus representations similarly across species and levels of stimulus complexity. Experiment 1 extends the generality of these past results by showing that similar shifts in generalization occur when individuals learn to discriminate (but not to identify) complex sounds. In addition, Spetch et al. speculated that discrimination training with multidimensional stimuli may be more likely to lead to shifts in generalization gradients that do not depend on asymmetries in the stimulus set used to test generalization. Our results are consistent with this hypothesis, but further experiments with complex stimuli are needed to identify how stimulus features impact learning-dependent generalization.

One stimulus feature that has rarely been incorporated into past studies of the peak shift effect is spatiotemporal variability. Bizo and McMahon (2007) recently reported that individuals trained to discriminate squares presented for different durations (e.g., 0.79 vs. 0.95 sec) showed a peak shift effect along the dimension of duration. Russell and Kirkpatrick (2007) acquired similar results with rats. Although the discriminandum in these experiments was time, the stimuli that they displayed were static. In contrast, our participants learned to distinguish temporally dynamic stimuli; the sounds they heard changed continuously. Comparable visual stimuli might include a shape that repeatedly moves around the screen in a figure-eight pattern or lights that flash at different frequencies (Sloane, 1964). Early work by Pavlov (1927) suggested that generalization mechanisms apply to variations in inputs over time in ways that parallel how they apply to spatial variations in inputs across sensory receptors. For example, Pavlov found similar patterns of generalization in dogs conditioned to salivate in the presence of a particular metronome rate and in dogs conditioned to respond to touches of a particular body part (see also Konorski, 1968). Experiments 1 and 2 provide further support for the supposition that training with temporally dynamic stimuli leads to generalization gradients comparable to those observed after training with static stimuli.

How is it possible that complex stimuli that vary over time could be generalized in the same way as simple stim-

uli that do not and that learning could affect both kinds of stimuli in similar ways? One possibility is that there are no “simple” static stimuli. For example, presenting a colored shape to an observer does not ensure that the inputs processed by that observer are simple or static. This will be determined by how the observer directs his or her gaze, head movements, and attention. In short, stimuli arise from a temporally dynamic stream of patterned receptor activation. Compared with visual reception, sound reception is less affected by both attention and body movements. Also, periodic sounds lead to predictable cortical responses (e.g., Orduña et al., 2005). From this perspective, stimuli with predictable temporal dynamics (e.g., FM sweeps) should generate the simplest stimulus representations (in the sense of being most consistent across trials). A better understanding of the stimulus representations that underlie discrimination learning and generalization may clarify when and how training shifts generalization gradients and why learning to distinguish stimuli varying along multiple dimensions might modulate the probability of generating a peak shift effect (Spetch et al., 2004).

Numerous past studies of learning-related shifts in generalization gradients have involved go/no-go training followed by testing with a set of novel stimuli centered on the reinforced stimulus (Purtle, 1973). This experimental design has led to a long history of debates about the role of inhibitory gradients and adaptation level adjustments as mechanisms for observed shifts (Ghirlanda & Enquist, 1999; MacLeod, Dodd, Sheard, Wilson, & Bibi, 2003; Purtle, 1973; Thomas, 1993; Thomas et al., 1991). In the present study, we controlled for differential reinforcement during training, as well as for asymmetries in stimulus sets during testing, and still found a learning-related shift in peak responding. Neither Spence’s (1937) nor Thomas et al.’s (1991; Thomas, Svinicki, & Vogt, 1973) explanation of the peak-shift effect are suitable accounts for these results. There are other less discussed explanations of learning-related shifts, but these also have explanatory troubles in this case. For example, a statistical decision theory interpretation might propose that during training, a distribution of internal events would develop for each training stimulus (Boneau & Cole, 1967). These distributions would overlap each other so that some internal events would be associated with responses to the S+, others would be associated with responses to the S–, and some would be associated with both (Boneau & Cole, 1967). An organism might respond in a way that maximizes its reward and minimizes its nonreward. To do so, it would set a criterion for a response that is not midway between the distributions, but shifted to minimize false alarms (responses in the presence of an S–). Such an explanation is consistent with findings that creatures will shift their criterion on the basis of how negative an S– is (Boneau & Cole, 1967; Heinemann, Avin, Sullivan, & Chase, 1969; Lynn et al., 2005). In our study, however, the participants still needed to make a response in order to be correct in the presence of an S–. Any shift in a criterion would have led to more misses of the S+ and less correct rejections of S–.<sup>4</sup>

Response competition offers another alternative. Here, it is proposed that during training, two gradients are developed for responses (Heinemann et al., 1969; Russell & Kirkpatrick, 2007). Peak shift occurs through a summation process similar to that proposed by Spence (1937). The advantage is that no need for inhibition is present, because the gradients, by being associated with different responses, subtract from each other to create shift (Russell & Kirkpatrick, 2007). Because the peaks of our gradients and others (Hanson, 1959; reviewed by Purtle, 1973) are not lower than those obtained after single stimulus training, response competition is not an optimal explanation. Alternatively, the competition between responses may cause participants’ stimulus representations to be in flux during training. This would be indicative of perceptual learning (Gibson, 1959; Goldstone, 1994; Hall, 1991; Liu, Mercado, Church, & Orduña, 2008; Saksida, 1999). In training that involves different consequences of different stimuli, participants could develop more distinct percepts of the sounds. Differentiating representations of the two stimuli would in turn make consequential differences between the training stimuli easier to separate. By being separated, however, the representations of these stimuli may become overlapped with representations of other stimuli on the dimension. Shift would then result, not from a summation of separate gradients, but from representational reorganization caused by the competition of responses.

The elemental model of discrimination learning (Blough, 1975; Ghirlanda & Enquist, 1999; McLaren & Mackintosh, 2002; Thorndike, 1932) is akin to this kind of process. It also appears to provide the simplest account of our findings, if it is assumed that the elements being analyzed include time-varying features. Previously employed techniques for measuring peak shift effects are insufficient for determining the nature of such a representational change (Petrov, Doshier, & Lu, 2005), partially because of the lack of an investigation into the temporal dynamics of learning-related shifts.

In many past studies of peak shift in humans, generalization tests occurred immediately after training, and no assessment was made of possible changes in gradients at different time points during or after training. The implicit assumption underlying this methodology is that the effects of training on generalization should be strongest soon after training. In fact, several models propose that generalization gradients are likely to flatten over time (Estes, 1997; Nairne, 1991; Riccio et al., 1992), which would mean that differences in gradients should decrease as the interval between training and testing increases. As was previously noted, however, there are reports of the peak shift effect persisting in nonhuman animals when days or weeks intervened between training and testing (Moye & Thomas, 1982; Spence, 1937; Thomas et al., 1960) and similar reports of stable generalization gradients over months and years (reviewed by Bouton et al., 1999). If similar mechanisms underlie peak shift in humans and other animals, one would expect that delayed testing should have little impact on the peak shift effect. The results of Experiment 2 confirm this prediction.

It remains unclear how long learning-induced shifts in generalization gradients might last; the effect was still observed after the longest interval tested in any species (3 weeks in pigeons; Thomas et al., 1960). Stimulus complexity, familiarity, naturalness, modality, or spatiotemporal variability might impact the time course of experience-dependent changes in generalization, and the quality and quantity of training that precedes the interval are also likely to be critical factors. An example of this can be seen in prior work showing that overtraining can actually cause the peak shift effect to dissipate (Gerry, 1971; Terrace, 1966). This raises the question of whether learning-related shifts reflect a useful end state or an intermediate/transitional state that is a prerequisite for a preferable end state. It could be the case that as learning progresses, the amount of peak shift changes quadratically rather than linearly. A novice would display no shift, an intermediate would show a peak shift, and an expert discriminator would lack this shift because of his enhanced ability to distinguish fine differences on the dimension. In this case, the expert's increased perceptual ability may lead to a decreased chance of being fooled by a novel stimulus.

Current explanations of learning-related shifts in generalization gradients often describe these phenomena as a side effect of discrimination learning. An alternative possibility that deserves further consideration is that gradient shifts are a direct effect of representational reorganization. Shifts in gradients that maximize differential responding to the stimuli most similar to the one experienced during training may increase the distinctiveness of the experienced stimulus more than simply increasing the response to it. This account could potentially explain why, in some cases, discrimination training not only shifts the peak during generalization but also raises it considerably for an unfamiliar stimulus (Hanson, 1959; Purtle, 1973).

In conclusion, future studies of learning-related shifts in generalization that control for (1) differential reinforcement during training, (2) asymmetries in stimulus sets during testing, (3) the temporal dynamics of stimuli, and (4) the time course of training and testing may provide useful data with which to further develop theories of stimulus generalization.

#### AUTHOR NOTE

This research was made possible by support from National Institute of Mental Health Grant MH67952 and National Science Foundation Science of Learning Center Grant SBE 0542013 to the Temporal Dynamics of Learning Center. We thank Estella H. Liu for her help in analysis and stimuli creation, and Laura Buckley, Loren Mallen, and Sara Rought for help carrying out the experiments. Thanks to David Smith and Peter Pfordresher for their guidance in developing the experimental design and data analyses and for providing useful comments on earlier versions of the manuscript. We also thank Jonathon Crystal and three anonymous reviewers for their helpful feedback on several versions of the manuscript. Correspondence concerning this article should be addressed to E. Mercado III, Department of Psychology, Park Hall, University at Buffalo, State University of New York, Buffalo, NY 14260 (e-mail: emiii@buffalo.edu).

#### REFERENCES

AHN, W.-K., & MEDIN, D. L. (1992). A two-stage model of category construction. *Cognitive Science*, *16*, 81-121.

- BIZO, L. A., & McMAHON, C. V. (2007). Temporal generalization and peak shift in humans. *Learning & Behavior*, *35*, 123-130.
- BLOUGH, D. S. (1975). Steady state data and a quantitative model of operant generalization and discrimination. *Journal of Experimental Psychology: Animal Behavior Processes*, *1*, 3-21.
- BONEAU, C. A., & COLE, J. L. (1967). Decision theory, the pigeon, and the psychophysical function. *Psychological Review*, *74*, 123-135.
- BOUTON, M. E., NELSON, J. B., & ROSAS, J. M. (1999). Stimulus generalization, context change, and forgetting. *Psychological Bulletin*, *125*, 171-186.
- CHENG, K. (1999). Spatial generalization in honeybees confirms Shepard's law. *Behavioural Processes*, *44*, 309-316.
- CHENG, K. (2000). Shepard's universal law supported by honeybees in spatial generalization. *Psychological Science*, *11*, 403-408.
- CHENG, K. (2002). Generalization: Mechanistic and functional explanations. *Animal Cognition*, *5*, 33-40.
- CHENG, K., & SPETCH, M. L. (2002). Spatial generalization and peak shift in humans. *Learning & Motivation*, *33*, 358-389.
- CHENG, K., SPETCH, M. L., & JOHNSTON, M. (1997). Spatial peak shift and generalization in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *23*, 469-481.
- DECHARMS, R. C., BLAKE, D. T., & MERZENICH, M. M. (1998). Optimizing sound features for cortical neurons. *Science*, *280*, 1439-1443.
- DICKINSON, J., & HEDGES, D. G. (1986). Adaptation level as an explanation of the peak shift in generalization with movement stimuli. *Journal of Motor Behavior*, *18*, 101-110.
- DOLL, T. J., & THOMAS, D. R. (1967). Effects of discrimination training on stimulus generalization for human subjects. *Journal of Experimental Psychology*, *75*, 508-512.
- DOUGHERTY, D. M., & LEWIS, P. (1991). Stimulus generalization, discrimination learning, and peak shift in horses. *Journal of the Experimental Analysis of Behavior*, *56*, 97-104.
- ESTES, W. K. (1997). Processes of memory loss, recovery, and distortion. *Psychological Review*, *104*, 148-169.
- FAGOT, J., GOLDSTEIN, J., DAVIDOFF, J., & PICKERING, A. (2006). Cross-species differences in color categorization. *Psychonomic Bulletin & Review*, *13*, 275-280.
- FORSTER, K. I., & FORSTER, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*, 116-124.
- GALIZIO, M. (1985). Human peak shift: Analysis of the effects of three-stimulus discrimination training. *Learning & Motivation*, *16*, 478-494.
- GALIZIO, M., & BARON, A. (1979). Human postdiscrimination gradients: The effects of three-stimulus discrimination training. *Animal Learning & Behavior*, *7*, 53-56.
- GERRY, J. E. (1971). Peak shift on tonal-frequency continuum: The effects of extinction and punishment. *Psychonomic Science*, *23*, 33-34.
- GHIRLANDA, S., & ENQUIST, M. (1999). The geometry of stimulus control. *Animal Behaviour*, *58*, 695-706.
- GHIRLANDA, S., & ENQUIST, M. (2003). A century of generalization. *Animal Behaviour*, *66*, 15-36.
- GIBSON, J. J. (1959). Perception as a function of stimulation. In S. Koch (Ed.), *Psychology: A study of a science* (Vol. 1, pp. 456-501). New York: McGraw-Hill.
- GOLDSTONE, R. (1994). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General*, *123*, 178-200.
- GUNZELMANN, G. (2008). Strategy generalization across orientation tasks: Testing a computational cognitive model. *Cognitive Science*, *32*, 835-861.
- GUTTMAN, N., & KALISH, H. I. (1956). Discriminability and stimulus generalization. *Journal of Experimental Psychology*, *51*, 79-88.
- HALL, G. (1991). *Perceptual and associative learning*. New York: Oxford University Press.
- HANSON, H. M. (1959). Effects of discrimination training on stimulus generalization. *Journal of Experimental Psychology*, *58*, 321-334.
- HEINEMANN, E. G., AVIN, E., SULLIVAN, M. A., & CHASE, S. (1969). Analysis of stimulus generalization with a psychophysical method. *Journal of Experimental Psychology*, *80*, 215-224.
- HELSON, H. (1964). *Adaptation level theory*. New York: Harper & Row.
- HELSON, H., & AVANT, L. L. (1967). Stimulus generalization as a function of contextual stimuli. *Journal of Experimental Psychology*, *73*, 565-567.

- HONIG, W. K., & STEWART, K. E. (1993). Relative numerosity as a dimension of stimulus control: The peak shift. *Animal Learning & Behavior*, **21**, 346-354.
- HONIG, W. K., & URCIOLI, P. J. (1981). The legacy of Guttman and Kalish (1956): 25 years of research on stimulus-generalization. *Journal of the Experimental Analysis of Behavior*, **36**, 405-445.
- IMAI, S., & GARNER, W. R. (1965). Discriminability and preference for attributes in free and constrained classification. *Journal of Experimental Psychology*, **69**, 596-608.
- JENKINS, H. M., & HARRISON, R. H. (1960). Effect of discrimination training on auditory generalization. *Journal of Experimental Psychology*, **59**, 246-253.
- JITSUMORI, M., & MAKINO, H. (2004). Recognition of static and dynamic images of depth-rotated human faces by pigeons. *Learning & Behavior*, **32**, 145-156.
- KONORSKI, J. (1968). *Conditioned reflexes and neuron organization*. New York: Hafner.
- LAZAREVA, O. F., WASSERMAN, E. A., & YOUNG, M. E. (2005). Transposition in pigeons: Reassessing Spence (1937) with multiple discrimination training. *Learning & Behavior*, **33**, 22-46.
- LEWIS, M. B., & JOHNSTON, R. A. (1999). Are caricatures special? Evidence of peak shift in face recognition. *European Journal of Cognitive Psychology*, **11**, 105-117.
- LIU, E. H., MERCADO, E., III, CHURCH, B. A., & ORDUÑA, I. (2008). The easy-to-hard effect in human (*Homo sapiens*) and rat (*Rattus norvegicus*) auditory identification. *Journal of Comparative Psychology*, **122**, 132-145.
- LYNN, S. K., CNAANI, J., & PAPA, D. R. (2005). Peak shift discrimination learning as a mechanism of signal evolution. *Evolution*, **59**, 1300-1305.
- MACLEOD, C. M., DODD, M. D., SHEARD, E. D., WILSON, D. E., & BIBI, U. (2003). In opposition to inhibition. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 43, pp. 163-214). San Diego: Academic Press.
- MCLAREN, I. P. L., BENNETT, C. H., GUTTMAN-NAHIR, T., KIM, K., & MACKINTOSH, N. J. (1995). Prototype effects and peak shift in categorization. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 662-673.
- MCLAREN, I. P. L., & MACKINTOSH, N. J. (2002). Associative learning and elemental representation: II. Generalization and discrimination. *Animal Learning & Behavior*, **30**, 177-200.
- MERCADO, E., III, ORDUÑA, I., & NOWAK, J. M. (2005). Auditory categorization of complex sounds by rats (*Rattus norvegicus*). *Journal of Comparative Psychology*, **119**, 90-98.
- MOYE, T. B., & THOMAS, D. R. (1982). Effects of memory reactivation treatments on postdiscrimination generalization performance in pigeons. *Animal Learning & Behavior*, **10**, 159-166.
- NAIRNE, J. S. (1991). Positional uncertainty in long-term memory. *Memory & Cognition*, **19**, 332-340.
- NEWLIN, R. J., RODGERS, J. P., & THOMAS, D. R. (1979). Two determinants of the peak shift in human voluntary stimulus generalization. *Perception & Psychophysics*, **25**, 478-486.
- ORDUÑA, I., MERCADO, E., III, GLUCK, M. A., & MERZENICH, M. M. (2005). Cortical responses in rats predict perceptual sensitivities to complex sounds. *Behavioral Neuroscience*, **119**, 256-264.
- PAVLOV, I. P. (1927). *Conditioned reflexes* (G. V. Anrep, Ed. & Trans.). London: Oxford University Press, Clarendon Press.
- PETROV, A. A., DOSHER, B. A., & LU, Z. L. (2005). The dynamics of perceptual learning: An incremental reweighting model. *Psychological Review*, **112**, 715-743.
- PURTLE, R. B. (1973). Peak shift: A review. *Psychological Bulletin*, **80**, 408-421.
- RICCIO, D. C., ACKIL, J. K., & BURCH-VERNON, A. (1992). Forgetting of stimulus attributes: Methodological implications for assessing associative phenomena. *Psychological Bulletin*, **112**, 433-445.
- RICCIO, D. C., URDA, M., & THOMAS, D. R. (1966). Stimulus control in pigeons based on proprioceptive stimuli from floor inclination. *Science*, **153**, 434-436.
- RODRIGUES, P. M., & MURRE, J. M. J. (2007). Rules-plus-exception tasks: A problem for exemplar models? *Psychonomic Bulletin & Review*, **14**, 640-646.
- RUSSELL, R., & KIRKPATRICK, K. (2007). The role of temporal generalization in a temporal discrimination task. *Behavioral Processes*, **74**, 115-125.
- SAKSIDA, L. M. (1999). Effects of similarity and experience on discrimination learning: A nonassociative connectionist model of perceptual learning. *Journal of Experimental Psychology: Animal Behavior Processes*, **25**, 308-323.
- SHEPARD, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, **237**, 1317-1323.
- SHEPARD, R. N., HOVLAND, C. I., & JENKINS, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs*, **75**, 42.
- SLOANE, H. N. (1964). Stimulus generalization along a light flicker rate continuum after discrimination training with several Ss. *Journal of the Experimental Analysis of Behavior*, **7**, 217-222.
- SPENCE, K. W. (1937). The differential response in animals to stimuli varying within a single dimension. *Psychological Review*, **44**, 430-444.
- SPETCH, M. L., & CHENG, K. (1998). A step function in pigeons' temporal generalization in the peak shift task. *Animal Learning & Behavior*, **26**, 103-118.
- SPETCH, M. L., CHENG, K., & CLIFFORD, C. W. G. (2004). Peak shift but not range effects in recognition of faces. *Learning & Motivation*, **35**, 221-241.
- TERRACE, H. S. (1966). Behavioral contrast and the peak shift: Effects of extended discrimination training. *Journal of the Experimental Analysis of Behavior*, **9**, 613-617.
- THOMAS, D. R. (1993). A model for adaptation-level effects on stimulus generalization. *Psychological Review*, **100**, 658-673.
- THOMAS, D. R., & BISTEY, G. (1964). Stimulus generalization as a function of the number and range of generalization test stimuli. *Journal of Experimental Psychology*, **68**, 599-602.
- THOMAS, D. R., & BURR, D. E. (1969). Stimulus generalization as a function of the delay between training and testing procedures: A reevaluation. *Journal of the Experimental Analysis of Behavior*, **12**, 105-109.
- THOMAS, D. R., & LYONS, J. (1968). Visual field dependency in pigeons. *Animal Behaviour*, **16**, 213-218.
- THOMAS, D. R., & MITCHELL, K. (1962). Instructions and stimulus categorizing in a measure of stimulus generalization. *Journal of the Experimental Analysis of Behavior*, **5**, 375-381.
- THOMAS, D. R., MOOD, K., MORRISON, S., & WIERTELAK, E. (1991). Peak shift revisited: A test of alternative interpretations. *Journal of Experimental Psychology: Animal Behavior Processes*, **17**, 130-140.
- THOMAS, D. R., OST, J., & THOMAS, D. (1960). Stimulus generalization as a function of the time between training and testing procedures. *Journal of the Experimental Analysis of Behavior*, **3**, 9-14.
- THOMAS, D. R., SVINICKI, M. D., & VOGT, J. (1973). Adaptation level as a factor in human discrimination learning and stimulus generalization. *Journal of Experimental Psychology*, **97**, 210-219.
- THOMAS, D. R., WINDELL, B. T., BAKKE, I., KREYE, S., KIMOSE, E., & APOSHYAN, H. (1985). Long-term memory in pigeons: I. The role of discrimination problem difficulty assessed by reacquisition measures: II. The role of stimulus modality assessed by generalization slope. *Learning & Motivation*, **16**, 464-477.
- THORNDIKE, E. (1932). *The fundamentals of learning*. New York: Teachers College Press.
- VERBEEK, E. L., SPETCH, M. L., CHENG, K., & CLIFFORD, C. W. G. (2006). Determinants of range effects in face recognition. *Learning & Behavior*, **34**, 229-240.
- WATTENMAKER, W. D. (1992). Relational properties and memory based category construction. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 1125-1138.
- WHITE, K. G., & THOMAS, D. R. (1979). Postdiscrimination stimulus generalization in humans: An extension of Galizio and Baron. *Animal Learning & Behavior*, **7**, 564-565.
- WILLS, S., & MACKINTOSH, N. J. (1998). Peak shift on an artificial dimension. *Quarterly Journal of Experimental Psychology*, **51B**, 1-31.
- WINTON, A. S., & BEALE, I. L. (1971). Peak shift in concurrent schedules. *Journal of the Experimental Analysis of Behavior*, **15**, 73-81.

## NOTES

1. .wav files of the stimuli used in both experiments are available as supplemental materials, downloadable from <http://lb.psychonomic-journals.org/content/supplemental>.

2. Separate analyses were done to determine whether more conservative post hoc tests and/or the inclusion of dropped participants would produce different results. Two-tailed  $t$  tests with Bonferroni corrections yielded identical results. We also found that when we included the data of the participants that were dropped, the statistical conclusions were the same.

3. In Experiment 2, separate analyses were done to determine whether more conservative post hoc tests would produce different results. Two-tailed  $t$  tests with Bonferroni corrections yielded identical significant results for all comparisons of conditions with a delay. However, the comparisons of the discrimination without delay with the identification without delay and the one-sample  $t$  test comparing the discrimination without delay group's gradient mean with the S+ dropped to marginal significance. These one-tailed comparisons were significant. These conditions are a direct replication of Experiment 1.

Increasing variability by including the dropped participants in the analysis of Experiment 2 also dropped the comparisons of the discrimination without delay with the identification without delay and the one-sample  $t$  test comparing the discrimination without delay group's gradient mean with the S+ below significance. All other main effects, interactions, and comparisons were identical to the results with these participants dropped.

4. In a separate analysis, the S+ response proportions from Experiments 1 and 2 were pooled together by condition (discrimination or identification). We contrasted the S+ response proportions to 7.9 Hz and the response proportions to 4.5, 5.2, 6.0, and 6.9 Hz as a group. The  $d'$  values were calculated and were as follows: discrimination,  $d' = 1.88$ ; identification,  $d' = 1.29$ . An independent-samples  $t$  test confirmed that these means were different [ $t(86) = 2.47, p = .016$ , Cohen's  $d = 0.525$ ]. This difference shows that the participants who discriminated in training developed a finer ability to distinguish the left side of the distribution from the S+ stimulus than did those who received identification training. There were no significant differences between groups on measures of  $c'$ .

#### SUPPLEMENTAL MATERIALS

The stimulus files from this study, in .wav format, are available as supplemental materials from <http://lb.psychonomic-journals.org/content/supplemental>.

(Manuscript received December 10, 2008;  
revision accepted for publication May 26, 2009.)