JASA ARTICLE

Perceptual, procedural, and task learning for an auditory temporal discrimination task[a)]

Leslie Q. Zhen[b)] and Sheila R. Pratt
Department of Communication Science and Disorders, University of Pittsburgh, Pittsburgh, Pennsylvania 15213, USA

ABSTRACT:
Perceptual learning reflects experience-driven improvements in the ability to detect changes in stimulus characteristics. The time course for perceptual learning overlaps with that for procedural learning (acquiring general skills and strategies) and task learning (learning the perceptual judgment specific to the task), making it difficult to isolate their individual effects. This study was conducted to examine the role of exposure to stimulus, procedure, and task information on learning for auditory temporal-interval discrimination. Eighty-three listeners completed five online sessions that required temporal-interval discrimination (target task). Before the initial session, listeners were differentially exposed to information about the target task's stimulus, procedure, or task characteristics. Learning occurred across sessions, but an exposure effect was not observed. Given the significant learning across sessions and variability within and across listeners, contributions from stimulus, procedure, and task exposure to overall learning cannot be discounted. These findings clarify the influence of experience on temporal perceptual learning and could inform designs of training paradigms that optimize perceptual improvements.
https://doi.org/10.1121/10.0017548

(Received 18 May 2022; revised 22 February 2023; accepted 23 February 2023; published online 20 March 2023)

[Editor: Laurie M. Heller] Pages: 1823–1835

I. INTRODUCTION

Perceptual learning reflects experience-driven improvements in the ability to detect changes in stimulus features, such as the amount of time separating a pair of sounds (Bratzke *et al.*, 2012; Delhommeau *et al.*, 2002; Hawkey *et al.*, 2004; Hussain *et al.*, 2012; Ortiz and Wright, 2009, 2010; Wright *et al.*, 1997). Demonstrations of perceptual learning often involve repetitive training with stimulus features, such as duration or frequency. In contrast, procedural learning reflects the process of acquiring general skills and strategies integral to the training experience and can include factors such as the experimental setting, user interface, and method of response (Delhommeau *et al.*, 2002; Hawkey *et al.*, 2004; Ortiz and Wright, 2009; Xu *et al.*, 2021). Task learning reflects mastering of the judgment specific to a task, such as discrimination of duration or frequency (Ahissar and Hochstein, 1997; Jeter *et al.*, 2009; Ortiz and Wright, 2009; Xu *et al.*, 2021).

Training protocols used to induce perceptual learning typically require repetitive, attention-directed practice with the feature of interest over many trials and sometimes across many days (Wright and Sabin, 2007; Wright and Zhang, 2009). Also, brief, single-session training protocols have been used to induce perceptual learning (Amitay *et al.*, 2006; Ortiz and Wright, 2009). Traditionally, perceptual learning is measured as the change in sensitivity to the feature of interest following a period of training (Wright and Zhang, 2009). In most instances, learning does not typically transfer to untrained stimuli (Karmarkar and Buonomano, 2003; Wright and Zhang, 2009), although some studies have demonstrated limited, asymmetric transfer of learning from trained to untrained stimuli (Amitay *et al.*, 2012; Bratzke *et al.*, 2012).

Many studies of perceptual learning have included pretraining to minimize the influences of procedural and task learning (Delhommeau *et al.*, 2002; Wright and Fitzgerald, 2001). This approach relies on general assumptions about the time courses of each type of learning. Steep gains early during learning are thought to denote procedural and task learning, whereas gradual improvements later in time reflect perceptual learning (Hawkey *et al.*, 2004; Hussain *et al.*, 2012; Ortiz and Wright, 2010; Wright and Fitzgerald, 2001; Wright and Sabin, 2007). However, these generalities lack specificity and may not correspond to all instances of sensory learning. In the time perception literature, the systematic investigation of these three types of learning are few, including in studies of auditory temporal-interval discrimination across multiple sessions. An alternative approach to isolating the different components of overall learning requires differentially exposing listeners to information about a target task of interest (Hawkey *et al.*, 2004; Ortiz and Wright, 2009, 2010). For example, Hawkey *et al.* (2004) measured procedural learning as the difference in performance on a frequency discrimination task (i.e., their target task) between groups that did and did not have prior training with the procedural aspects of their target task. Furthermore, perceptual learning was measured as the difference in frequency selectivity between groups that did and

[a)]Portions of this work were presented in "Differentiating perceptual, procedural, and task learning for an auditory temporal discrimination task," 181st meeting of the Acoustical Society of America, Seattle, WA, USA, November 2021.
[b)]Electronic mail: lez35@pitt.edu

did not have prior training with the perceptual task. Of note, the *perceptual learning* in Hawkey *et al*. (2004) is a combination of the perceptual and task learning in this study. This approach to isolate the different aspects of overall learning circumvents assumptions about the time courses of each type of learning.

A. Stimulus experience and learning

Exposure to a stimulus can improve sensitivity to that stimulus. Many studies support a hypothetical threshold that related neural processes involved in learning must surpass for perceptual learning to occur (Seitz and Dinse, 2007; Szpiro *et al*., 2014; Wiggs and Martin, 1998; Wright *et al*., 2010; Wright and Sabin, 2007). Active practice-related stimulus exposures foster *permissive signals* that drive learning on the target perceptual task. Further stimulation enables these related neural processes to surpass the hypothetical threshold of learning. For example, Wright *et al*. (2010) observed enhanced perceptual learning for a frequency discrimination task when blocks were interleaved with blocks of auditory temporal-interval discrimination compared to frequency discrimination training alone. Follow-up experiments revealed that this enhancement to frequency perceptual learning was observed regardless of whether the interval task exclusively preceded or succeeded training on the target frequency discrimination task. Enhancements to frequency discrimination performance at the posttest were attributed to additional trials of stimulus exposures that were interleaved with the frequency discrimination task. They postulated that target frequency discrimination performance fostered permissive signals that placed related neural processes involved in learning in a highly malleable, sensitized state. The trials of stimulus exposures encountered outside of the target frequency discrimination performance provided the additional stimulation needed to surpass the hypothetical threshold for perceptual learning. Notably, perceptual enhancements were observed even when additional sensory stimulation encountered independent of the target frequency discrimination performance was behaviorally irrelevant to the trained nontarget interval task. This has interesting implications for attention, which has been traditionally regarded as requisite for perceptual learning. Given that enhancements to frequency discrimination were observed even though listeners were instructed to attend to duration, not frequency, during the interval task, findings suggest that attention might have played a mediating role for perceptual learning.

Indeed, many studies observed that exposure to a task-irrelevant feature during training can enhance performance on a subsequent target task that evaluates the same feature, and this perceptual learning can occur even without attention and awareness (Meuwese *et al*., 2013; Seitz and Dinse, 2007; Seitz *et al*., 2009; Seitz *et al*., 2010; Seitz and Watanabe, 2003, 2005, 2009; Vlahou *et al*., 2012; Watanabe *et al*., 2001). This phenomenon has been coined *task-irrelevant perceptual learning*. One of the earliest demonstrations of task-irrelevant perceptual learning comes from Watanabe *et al*. (2001), who exposed participants to a background that contained subthreshold levels of visual coherent motion while they actively trained on a letter identification task. The background coherent motion was deemed task-irrelevant as it was irrelevant to the letter identification task that participants were asked to complete. Exposure to this subthreshold coherent motion improved subsequent performance on suprathreshold direction indication, coherent motion detection, and direction discrimination tasks for the exposed motion direction. They proposed that frequency, not salience, of stimulus exposure sensitized the perceptual system to that feature. In audition, corroborating evidence for this hypothesis comes from findings that training with at least 360 trials per day on auditory temporal-interval discrimination is required to elicit perceptual learning (Wright and Sabin, 2007).

Follow-up investigations suggest that task-irrelevant perceptual learning is conditional on attention-based requirements of the trained task (Ahissar and Hochstein, 1993; Bruns and Watanabe, 2019; Leclercq and Seitz, 2012; Seitz and Watanabe, 2008). Namely, the perceptual system's interaction with task-irrelevant stimuli is enhanced when the task-irrelevant stimuli are presented during behaviorally relevant events (Seitz and Watanabe, 2003, 2009; Leclercq and Seitz, 2012). For example, Seitz and Watanabe (2003) observed enhanced performance on a suprathreshold random-dot motion detection task following exposure to background motion direction in a letter identification task when the motion direction was paired with the task's targets but not for other background motion directions paired with distractors. They conjectured that the perceptual system reinforced an internal reward signal on recognizing the pairing of the motion direction and task's targets. This internal reward signal then gated learning on the target random-dot motion detection task for the paired direction (Roelfsema *et al*., 2010; Seitz and Dinse, 2007; Seitz *et al*., 2009; Seitz and Watanabe, 2005; Watanabe and Sasaki, 2015). This learning signal is thought to diffusely enhance representations for any stimuli, task-relevant or task-irrelevant, that are systematically and temporally paired with the task targets (Seitz *et al*., 2005).

Although findings of robust task-irrelevant perceptual learning typically use paradigms where the task-irrelevant stimuli are presented at subthreshold or perithreshold levels (see Leclercq and Seitz, 2012, for a demonstration of task-irrelevant perceptual learning using suprathreshold stimuli), many studies adopting gamified training approaches reported learning of task-irrelevant stimuli presented at suprathreshold levels (Amitay *et al*., 2006; Lim and Holt, 2011; Gabay *et al*., 2015; Vlahou *et al*., 2012; Vlahou *et al*., 2019; Wade and Holt, 2005; Wiener *et al*., 2019). For example, Wade and Holt (2005) trained participants to play a 30-min videogame that involved capturing two friendly "alien" characters and destroying two enemy aliens. Suprathreshold, nonspeech exemplars drawn from one of four categories were consistently and temporally paired with the same alien character, each of which had unique shapes, colors, and movement trajectories. Learning the nonspeech categories became increasingly beneficial for game performance as difficulty progressed across levels (e.g., blocks), although no information was provided about the existence and behavioral relevance of the categories. Posttest accuracy was significantly above chance on an explicit categorization task requiring identification of alien characters based on trained and novel sound exemplars. Further, posttest categorization accuracy was positively correlated with several metrics of in-game performance (e.g., mean high score attained, mean high level reached, mean line-of-sight distance). Wade and Holt (2005) concluded that the audiovisual correlation patterns imparted through gameplay promoted learning of the nonspeech categories following incidental exposure to category exemplars, although this learning was observed only for categories characterized by structured onset variability patterns in higher dimensional acoustic space that consisted of onset trajectory and steady-state frequency cues (see Fig. 4 in Wade and Holt, 2005). Further, the information acquired differed between incidental and explicit training procedures. Specifically, unlike incidental training via gameplay, training on an explicit, unsupervised categorization task produced categorization patterns that reflected greater difficulty for categories defined by dynamic spectrotemporal changes of the second spectral peak at stimulus onset than those defined by these changes at the offset. The videogame used in Wade and Holt (2005) has been adapted to demonstrate incidental learning of nonnative speech categories (Lim and Holt, 2011).

The incidental auditory category learning demonstrated in Wade and Holt (2005) and Lim and Holt (2011) might not be entirely passive because auditory categories consistently predicted visual location and the associated motor response. Thus, category learning might require learning these associations beyond passive exposure alone. Gabay *et al*. (2015) evaluated the possible drivers of incidental category learning using a simplified task. Participants trained on five blocks of visual target location detection. During every trial, five repetitions of auditory exemplars identical to those used in Wade and Holt (2005) predicted visual target location and the associated motor response. Certain aspects of the exemplar-to-location mappings used for blocks 1–3 were modified in block 4 and restored in block 5. Disrupting the structured category-to-location mappings in block 4 slowed responses to target detection, suggesting incidental learning of auditory categories. Target detection responses were not slowed when trained exemplars were replaced with novel category-consistent exemplars during block 4, which otherwise maintained the exemplar-to-location mappings from blocks 1–3, proposing incidental learning of auditory categories and learning transfer to novel exemplars. In a different condition, detection responses were not slowed when the exemplar-to-location mappings from blocks 1–3 were replaced with novel mappings that disregarded category structure (e.g., arbitrary exemplar-to-location mappings) during block 4, suggesting that simple memorization of sound-to-location associations were not the primary driver of incidental category learning. The amount of exemplar exposures was identical between the condition with arbitrary exemplar-to-location mappings and the condition with structured category-to-location mappings. Despite this, posttest accuracy on a surprise overt category labelling task was at chance for the former condition and significantly above chance for the latter, recommending that passive exposure alone cannot account for incidental auditory category learning. Gabay *et al*. (2015) proposed that the consistent pairings of visual (e.g., visual target) and motor (e.g., task-related motor response) events with auditory exemplars drove learning of auditory categories. This explanation is reminiscent of the reinforcement learning accounts used to explain task-irrelevant perceptual learning (Seitz and Watanabe, 2003, 2009).

Despite the evidence against passive exposure-based accounts of learning task-irrelevant, suprathreshold stimuli in the auditory category learning literature, there has been demonstrations of passive learning of task-irrelevant auditory stimuli. Amitay *et al*. (2006) observed enhanced frequency discrimination after playing a visuospatial Tetris game (SodaCan, http://sivut.koti.soon.fi/sodacan) while passively listening to playback of $\geq$400 trials of another participant's frequency discrimination performance adaptively tracking 75% correct. Of note, participants were instructed to ignore the auditory stimuli, which were behaviorally irrelevant to the Tetris game. Enhancements were greater than that observed for another group of participants that played the Tetris game in silence, although both groups demonstrated significantly greater learning than those without any training. Improvements in frequency discrimination were thought to reflect stimulus-driven enhancements for those who passively listened to auditory stimuli and procedure-driven enhancements for those who played Tetris silently.

Overall, these findings provide evidence that subthreshold and suprathreshold stimulus exposures encountered outside of the target task can enhance learning. Attention does not appear to be requisite to contribute to perceptual enhancements in some instances. These stimulus exposure approaches were leveraged to enhance perceptual learning during this study's target task.

B. Procedure experience and learning

Experience with a task's procedures may enhance performance on perceptual tasks with the same or similar procedures. Although few studies have formally investigated the role of procedural learning for temporal information, some suggest that learning does not culminate solely from feature learning (Ahissar and Hochstein, 2004; Ahissar *et al*., 2009; Cohen *et al*., 2013; Hochstein and Ahissar, 2002; Xu *et al*., 2021). For example, Ortiz and Wright (2009) demonstrated that exposure to an interaural timing difference discrimination task's stimulus and procedures contributed to overall learning of the task. It is unknown whether these results extend to temporal perceptual learning.

Task structure, which includes factors or strategies that influence the decision-making process, has been found to affect temporal perceptual learning (Cohen *et al.*, 2013; Xu *et al.*, 2021). One interesting demonstration comes from Xu *et al.* (2021), where they provided evidence that temporal perceptual learning is influenced by the trained stimulus features and task structure. Although Xu *et al.* (2021) did not distinguish between the procedure- and task-related aspects of temporal learning, their use of task structure is a combination of this study's definitions of procedural and task learning. In their experiments, participants trained on an auditory temporal-interval discrimination task that consisted of either a fixed set of eight possible comparison intervals (fixed-interval group) or comparison intervals that were randomly selected from a Gaussian distribution (random-interval group). They observed significant reductions (improvements) in duration thresholds in the fixed-interval group but no gains in the random-interval group after five training sessions. This difference was attributed to the fact that the fixed-interval group, unlike the random-interval group, had access to comparison intervals that were more predictable, which enhanced strategy learning related to the statistical regularity of the comparison intervals, and contributed to overall learning.

In the same experiment, Xu *et al.* (2021) provided evidence for task learning. The interval task used during the five training sessions was a bisection task in which the participant indicated whether comparison intervals were longer or shorter than a standard interval that was presented once at the start of each session. Prior to and after the bisection task, participants completed a different variation of temporal-interval discrimination that was a two-interval comparison task during which the participant indicated which of two intervals was shorter. Although learning was observed for the fixed-interval group on the trained bisection task, this learning did not transfer to the comparison task during the testing phase. The task specificity of learning was attributed to learning response strategies during the trained bisection task that differed from the response strategies needed for the transfer comparison task.

C. Task experience and learning

As demonstrated in Xu *et al.* (2021) and by others, experience with a task may influence later performance on the same task. For example, studies of temporal-interval discrimination have demonstrated learning transfer across frequencies (Karmarkar and Buonomano, 2003; Wright *et al.*, 1997) and modalities from audition to vision (Bratzke *et al.*, 2012; Bratzke *et al.*, 2014; McGovern *et al.*, 2016). Key to these findings is that most demonstrate learning transfer from a trained task to the same task but not an untrained task, indicating that something about the acquired information is specific to the trained task. Indeed, a characteristic observation of repetitive training protocols is that improvements on trained tasks fail to transfer to untrained tasks even when the target stimulus is identical between tasks (Amitay *et al.*, 2006; Delhommeau *et al.*, 2002; Liu and Weinshall, 2000; Xu *et al.*, 2021). Thus, learning transfer of task information is expected only across identical tasks.

Ortiz and Wright (2009) suggested that procedural and perceptual learning but not task learning contributes to the learning process. In their study, listeners trained on one of three discrimination tasks (temporal-interval, interaural level difference, and interaural timing difference) one day before training on a target interaural timing difference discrimination task. Listeners who trained on interaural level difference were exposed to the target task's procedure and lateralization task. These listeners performed similarly to interval-trained listeners who were exposed only to the target task's procedure despite the fact that both groups performed better than controls naive to the target task. They concluded that task learning did not contribute toward overall learning for interaural timing difference discrimination but procedural learning did in that the interval-trained listeners performed better than untrained, naive controls. Similarly, listeners who trained on the interaural timing difference were exposed to the target task's stimulus, procedure, and task. These listeners performed better than those who trained on interaural level difference, suggesting that perceptual learning contributed toward overall learning. It is unknown whether these results extend to temporal-interval discrimination.

II. RESEARCH QUESTIONS AND HYPOTHESES

The goal of this study was to assess the extent to which exposure to the stimulus, procedure, and/or task of an auditory temporal-interval discrimination task (target task) influences learning on the target task. Before training on the target task, participants trained on one of three exposure tasks: (a) auditory temporal-interval discrimination (interval exposure-trained), (b) frequency discrimination with timing information [(FDT)-trained], and (c) frequency discrimination without timing information [(FD)-trained]. Control participants naive to the target task also were recruited. Hereafter, we use *interval exposure task* to refer to the one session of temporal-interval discrimination completed during the exposure phase and *interval target task* to refer to the five sessions completed throughout the training phase (see Fig. 1).

Table I shows the shared elements between the exposure and target tasks. For interval exposure-trained participants, gains on interval target task performance beyond those for FDT-trained participants denote task learning. For FDT-trained participants, improvements in interval target task performance beyond those for FD-trained participants imply perceptual learning. Last, for FD-trained participants, gains on interval target task performance beyond those for naive controls suggest procedural learning. Ortiz and Wright (2009) reported that perceptual and procedural learning but not task learning contributes to overall learning for the interaural timing difference. If the learning patterns for temporal-interval discrimination are consistent with those for interaural timing difference, we expected that the FDT-trained listeners

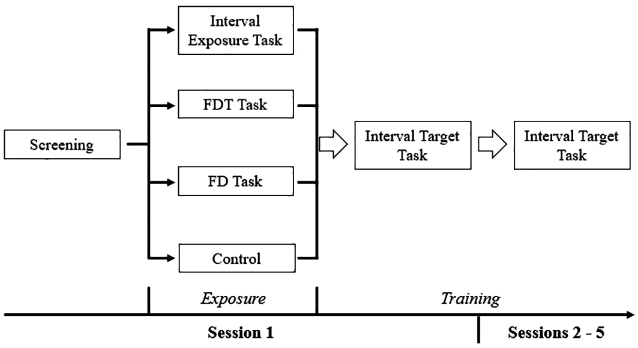

FIG. 1. Study design. Interval exposure task = interval target task = auditory temporal-interval discrimination and control = no exposure.

would improve more than the FD-trained listeners, the FD-trained listeners would improve more than the naive control group, and the interval exposure- and FDT-trained listeners would perform similarly on the interval target task. That said, consistent with evidence that task learning contributes to overall temporal learning (Xu *et al.*, 2021), interval exposure-trained listeners were expected to improve more than the FDT-trained listeners on the interval target task.

III. METHODS AND PROCEDURES

A. Participants

Eighty-three participants (60 female, 23 male) aged 18 to 30 years old [$M = 23.4$, standard deviation (SD) $= 3.0$] qualified for the study. The G*Power version 3.1.9.2 (Faul *et al.*, 2007) software was used to estimate the minimum sample size required to test an exposure-by-session interaction. A sample size of $n = 48$ was required to achieve 80% power for detecting an effect of 0.2 at a significance criterion of $\alpha = 0.05$ and correlation among repeated measures of 0.5 for a 4×5 (group $\times$ session) mixed design analysis of variance (ANOVA). After considering the entirely online procedure of this study, an effect size of 0.2 was selected to be a slightly more conservative estimate than those reported in relevant studies of auditory perceptual learning conducted in controlled laboratory settings (Ortiz and Wright, 2009, 2010; Xu *et al.*, 2021). A correlation among repeated measures of 0.5 was selected as it is a relatively moderate estimate for multisession auditory perceptual learning and given the trajectories reported in relevant studies (Bratzke *et al.*, 2012; Bratzke *et al.*, 2014; Cohen *et al.*, 2013; Karmarkar and Buonomano, 2003; Wright *et al.*, 1997; Xu *et al.*, 2021). The nonsphericity correction coefficient was set to one to assume sphericity.

TABLE I. Shared elements between the exposure and target tasks. Interval exposure = auditory temporal-interval discrimination and control = no exposure.

	Interval target task elements		
Exposure tasks	Stimulus	Procedure	Task
Interval exposure	Yes	Yes	Yes
FDT	Yes	Yes	No
FD	No	Yes	No
Control	No	No	No

TABLE II. Participant demographics.

		Exposure			
Factor		Interval	FDT	FD	Control
N		23	18	22	20
Sex	Female	12	15	16	15
	Male	11	3	6	5
Education[a]	<High school	1	0	0	0
	High school	8	7	7	7
	Associates	0	0	1	0
	Bachelors	11	8	12	10
	Masters	3	2	2	3
	Doctoral	0	1	0	0
		M (SD)			
Age		22.8 (2.7)	23.9 (3.7)	23.3 (2.7)	23.6 (2.8)
4-WTA[b]	Right ear	14.0 (6.6)	13.7 (6.8)	14.1 (8.1)	13.2 (7.7)
	Left ear	13.9 (7.0)	11.7 (6.6)	14.3 (8.2)	11.8 (8.3)

[a]Highest level of education completed.
[b]Warble tone average, calculated as the mean of 0.5, 1.0, 2.0, and 4.0 kHz dB hearing level (HL).

Given an expected completion rate of 70%, the minimum recruitment target was set to $n = 72$ (48/0.70, rounded up to the nearest whole number divisible by the number of groups) with $n = 18$ participants in each group. Complications related to the entirely online (necessitated by the COVID-19 pandemic) and longitudinal nature of this study resulted in more data loss than expected, and several participants performed aberrantly (see Sec. III E), hence, additional participants were enrolled.

Table II summarizes the demographic information for all of the qualified participants. The participants had normal or near-normal hearing, normal or corrected-to-normal vision for both eyes, and reported no speech, language, attention, and neurological impairments. All of the participants were naive to this and similar psychophysical experiments. Participants were recruited through the Pitt+Me research registry at the University of Pittsburgh and various Facebook (Meta Platforms, Inc., Menlo Park, CA) groups frequented by the target population and compensated for their time. All of the procedures were approved by the University of Pittsburgh's Institutional Review Board and informed consent was obtained from all of the participants.

B. Screening tasks

Hearing thresholds were screened using an online testing application.[1] Calibration was achieved following the site's level matching procedure (see Sec. 1 of the online testing application[1]), which required participants to adjust their computer's audio output level to match the levels of sounds of hands rubbing together, which were produced binaurally from the headphones and by the participants themselves after having removed the headphones from their ears. Hearing thresholds were determined using warble tones rather than conventional pure tones to minimize resonance from the participant's headphones and environment. Each ear was tested separately. Normal hearing was defined as hearing thresholds of ≤ 35 dB HL for 0.25, 0.5, 1, 2, 3, 4, 6, and 8 kHz (American Speech-Language-Hearing Association, 2005) in both ears.

Visual acuity was evaluated online with the application available.[2] The participants were instructed to indicate the direction of the letter *E*, which got progressively smaller when viewed from a constant distance. Each eye was tested separately.

Demographic and other eligibility information were screened using a questionnaire, andinformation collected included history of speech, language, hearing, attention, and neurological impairments. Also, the participants were asked to describe past participation in research studies if there was any.

C. Apparatus

The study was conducted online in the participants' natural environment because of restrictions related to the COVID-19 pandemic. The participants were required to have access to a computer, smartphone with Apple's iOS (Apple Inc., Cupertino, CA) or Google's Android (Alphabet Inc., Mountain View, CA) operating systems, wired circumaural headphones, and a webcam. Stimuli were presented using Pavlovia version 2020.2, which did not require an internet connection during experimental testing (Peirce *et al.*, 2019).

1. Stimuli

The stimuli consisted of an empty acoustic interval bounded by pairs of identical 15 ms, 1 kHz tones with 5 ms raised-cosine rise/fall ramps. The tone pairs were presented sequentially with a fixed 100 ms standard intertone interval between the first pair of tones and 1000 ms separation between the first and second pairs. The intertone interval for the second pair was variable. The stimuli were presented binaurally via circumaural headphones at a comfortable and audible level, which was estimated at the beginning of each session using white noise [10 s duration, 0.9–1.1 kHz passband, 44.1 kHz sampling rate, root mean square (RMS) = –20.3 dB] calibration to 70 dBZ employing either the NIOSH Sound Level Meter application for iOS (NIOSH, 2020) or the Decibel X–dB Sound Level Meter, Noise Detector application for Google's Android (SkyPaw Co., Ltd., 2021) and verified through verbal confirmation from the participant.

2. Interval target task

The interval target task was an auditory temporal-interval discrimination task adapted from Bratzke *et al.* (2012). The participants were instructed to indicate whether the second of two pairs of tones was separated by a shorter or longer duration than the first pair by using the left and right arrow keys on their keyboard. Figure 2 illustrates one trial of the interval target task. As indicated above, the intertone interval for the first pair of tones was fixed at 100 ms (the standard interval). The intertone interval for the second pair of tones varied across trials and included one of eight possible durations: 79, 85, 91, 97, 103, 109, 115, and 121 ms (the comparison interval). The standard was always presented first, and the comparison interval was selected randomly with replacement until each of the 8 possible durations was presented 20 times per block. The interval separating each trial was selected randomly from one of four possible durations: 800, 900, 1000, and 1100 ms.

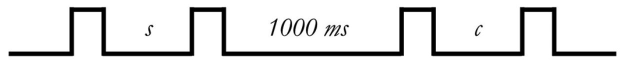

FIG. 2. One trial of the auditory tasks. A trial consisted of four auditory events and three primary intervals of silence. For the auditory temporal-interval discrimination task, each of the four events on the line represented a 1 kHz tone pip that was 15 ms in duration with 5 ms raised-cosine rise/fall ramps. *s* is the standard interval, which was fixed at 100 ms; *c* is the comparison interval, which was one of eight possible durations: 79, 85, 91, 97, 103, 109, 115, and 121 ms. Trials for the two variants of the frequency discrimination tasks shared the same format, except that the second pair of tone pips varied in frequency across trials based on one of eight possible frequencies: 979, 985, 991, 997, 1003, 1009, 1015, and 1021 Hz. Also, *c* was fixed at 100 ms for the frequency discrimination task without timing information.

3. Frequency discrimination with timing information

A FDT task required participants to indicate whether the second pair of tones was lower or higher in frequency than the first pair. Trials were identical to those for the interval target task (see Fig. 2), except that the second pair of tones varied by frequency across trials. Tones in the same pair were equal in frequency. The standard frequency was fixed at 1000 Hz. The comparison frequency was one of eight possible frequencies: 979, 985, 991, 997, 1003, 1009, 1015, and 1021 Hz. Each comparison frequency always co-occurred with the same intertone comparison interval. For example, a stimulus pair with a 979 Hz comparison frequency always had a 79 ms intertone interval and a 985 Hz pair had an 85 ms interval. Although coupling frequencies with durations may provide a temporal cue that enhances FDT performance beyond that for FD alone, cue-related improvements to frequency discrimination were orthogonal to our intent to use the FDT and FD tasks to expose participants to elements of the interval target task and not to evaluate frequency selectivity.

4. Frequency discrimination without timing information

The FD task was identical to the FDT task except that the intertone comparison interval was fixed at 100 ms.

D. Procedure

Figure 1 illustrates the study design. The experiment consisted of five sessions. On session (day) 1, participants were assigned randomly to train on an interval, FDT task, or FD task immediately prior to interval target task training. A control group did not receive any training before starting the interval target task. After this exposure phase, all of the participants completed one session of interval target task training. During sessions 2–5, participants trained only on the interval target task. For all of the tasks, each session consisted of 4 blocks of 160 trials. Visual feedback was provided after each trial. The maximum time between sessions was limited to 3 days.

E. Statistical analyses

The primary outcome was the number of *long* responses for each of the eight comparison intervals for the interval target task. Logistic curves were fitted to data of the eight comparison intervals as a function of probability of responding *long* using maximum likelihood for each participant and session without bootstrapping while estimating psychophysical thresholds. Psychometric functions were fit using the QuickPsy package in *R* (see Fig. 3; Linares and López-Moliner, 2016). For each participant, group, and session, the difference limen (DL) was estimated to be

$$\mathrm{DL} = \frac{x_{0.75} - x_{0.25}}{2}, \tag{1}$$

where $x_{0.25}$ and $x_{0.75}$ indicate the durations when the participant responded long 25% or 75% of the time, respectively. A mixed ANOVA was used to estimate the extent to which exposure to elements of the interval target task affected interval target task duration DLs across sessions. The same analyses were repeated on the slopes of the fitted functions.

Prior to analyses, outliers were identified and removed mirroring procedures used in relevant studies of auditory perceptual learning (Bratzke *et al.*, 2014; Ortiz and Wright, 2009, 2010; Xu *et al.*, 2021). First, interval target task data were prescreened to remove participants with completely flat psychometric functions (interval = 3, FDT = 1, FD = 4, and control = 2). Then, data were excluded for participants with deviant duration DL estimates, defined as >1.5 times the interquartile range and >25 ms. An upper cutoff of 25 ms was selected because it was slightly larger than the upper range of data from studies using similar interval tasks (Bratzke *et al.*, 2012; Bratzke *et al.*, 2014; Xu *et al.*, 2021) and after considering the individual trajectories for participants to maximize conservativeness of outlier removal given the large observed variances. Duration DLs from all of the participants were pooled to determine session 1 interval target task outliers. Then, duration DLs were analyzed by group to determine the outliersin sessions 2–5. In summary, data exist for at least 1 interval target task training session for 20 of 23 interval exposure-trained, 17 of 18 FDT-trained, 18 of 22 FD-trained, and 18 of 20 naive control participants. Incomplete data for one or more sessions resulted from attrition (FDT = 1, FD = 1, and control = 2) and technological difficulties (interval exposure = 4 and control = 2).

For the exposure tasks, outlier detection was performed for each group separately but otherwise followed the same procedure as outlined above. The exposure task results reported reflect data from 18 interval exposure-trained, 15 FDT-trained, and 16 FD-trained listeners.

The script for analysis and de-identified datasets are available online.[3]

IV. RESULTS

A. Effect of interval exposure on the interval target task

Figures 4 and 5 show the changes in duration DLs and slopes across sessions for each group and participant, respectively. A mixed ANOVA was used to investigate the effect of exposure on duration DLs across the five training sessions. The duration DLs were submitted to a 4 × 5 mixed ANOVA with group (interval, FDT, FD, and control) as the between-subjects factor and session (1, 2, 3, 4, and 5) as the within-subjects factor. The interaction between group and session was not statistically significant, $F(12,212) = 0.447$,

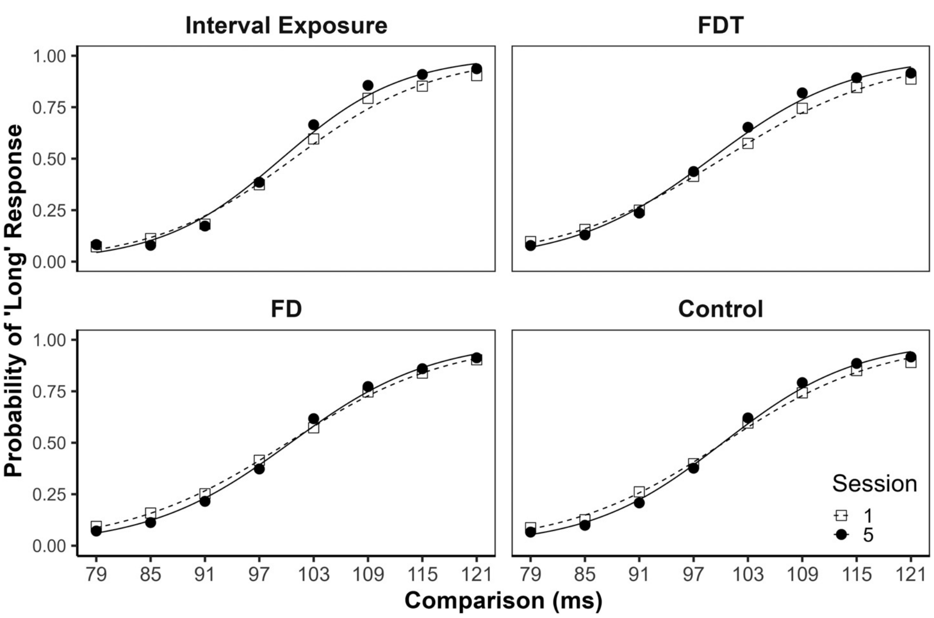

FIG. 3. Change in psychometric functions from initial to final sessions of target temporal-interval discrimination training. Interval exposure = auditory temporal-interval discrimination ($n_{\mathrm{sessions}1,5}$ = 18, 18); FDT ($n_{\mathrm{sessions}1,5}$ = 14,16); FD ($n_{\mathrm{sessions}1,5}$ = 16,17); control = no exposure ($n_{\mathrm{sessions}1,5}$ = 15,17). For clarity, only sessions 1 and 5 are plotted.

https://doi.org/10.1121/10.0017548

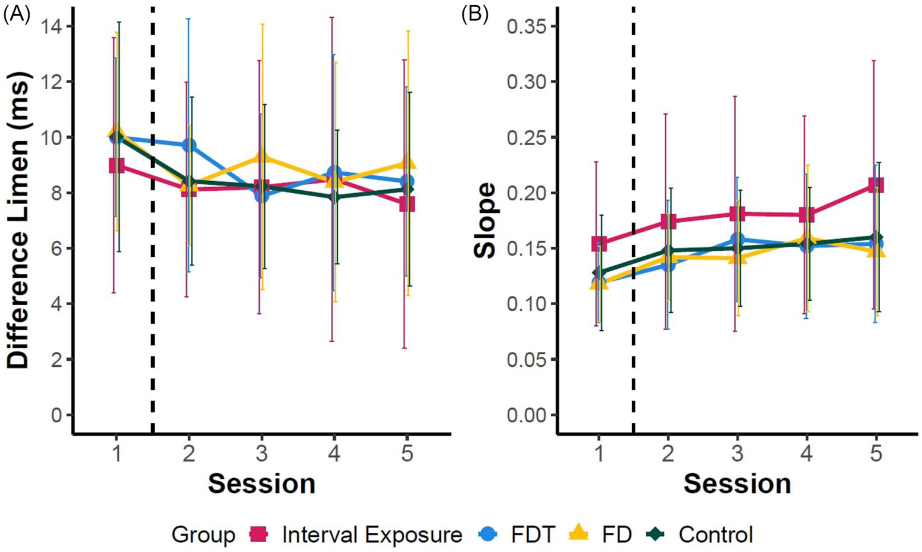

FIG. 4. (Color online) Change in (A) mean DLs and (B) slopes for temporal-interval discrimination across sessions. Interval exposure = auditory temporal-interval discrimination (n_{total} = 20); FDT (n_{total} = 17); FD (n_{total} = 18); control = no exposure (n_{total} = 18). The vertical dashed line separates interval-training during exposure from interval target task training. The error bars represent SD.

$p = 0.942$, and $\eta^2_p = 0.025$. No between group differences were observed, $F(3,53) = 0.159$, $p = 0.923$, and $\eta^2_p = 0.009$. Despite this, duration DLs decreased across sessions, $F(4,212) = 14.013$, $p = 3.67 \times 10^{-10}$, and $\eta^2_p = 0.209$, suggesting significant learning over time. Analysis on the slopes returned identical conclusions, where there was no interaction effect, $F(12,212) = 0.710$, $p = 0.741$, and $\eta^2_p = 0.039$; no effect of group, $F(3,53) = 1.181$, $p = 0.326$, $\eta^2_p = 0.063$; but a statistically significant steepening of slopes across sessions, $F(4,212) = 21.776$, $p = 4.55 \times 10^{-15}$, and $\eta^2_p = 0.291$.

Figure 6 shows additional analyses of the duration DLs and slopes that consider the maximal improvement in temporal-interval discrimination for each group. These analyses were motivated by the possibility that factors such as fatigue, inattention, and lack of motivation coupled with the entirely online administration might have contributed to the large observed variances (see Fig. 4) that masked possible exposure effects. A mixed ANOVA comparing the duration DLs from the initial to best sessions revealed a significant effect of session, $F(1,57) = 102.530$, $p = 2.37 \times 10^{-14}$, $\eta^2_p = 0.643$, but not group, $F(3,57) = 0.438$, $p = 0.727$, $\eta^2_p = 0.023$. The interaction was not significant, $F(3,57) = 0.119$, $p = 0.949$, $\eta^2_p = 0.006$. Comparisons of the duration DLs from the initial to final sessions revealed a significant effect of session, $F(1,54) = 23.131$, $p = 1.25 \times 10^{-5}$, $\eta^2_p = 0.300$, but not group, $F(3,54) = 0.424$, $p = 0.737$, $\eta^2_p = 0.023$, and interaction, $F(3,54) = 0.143$,

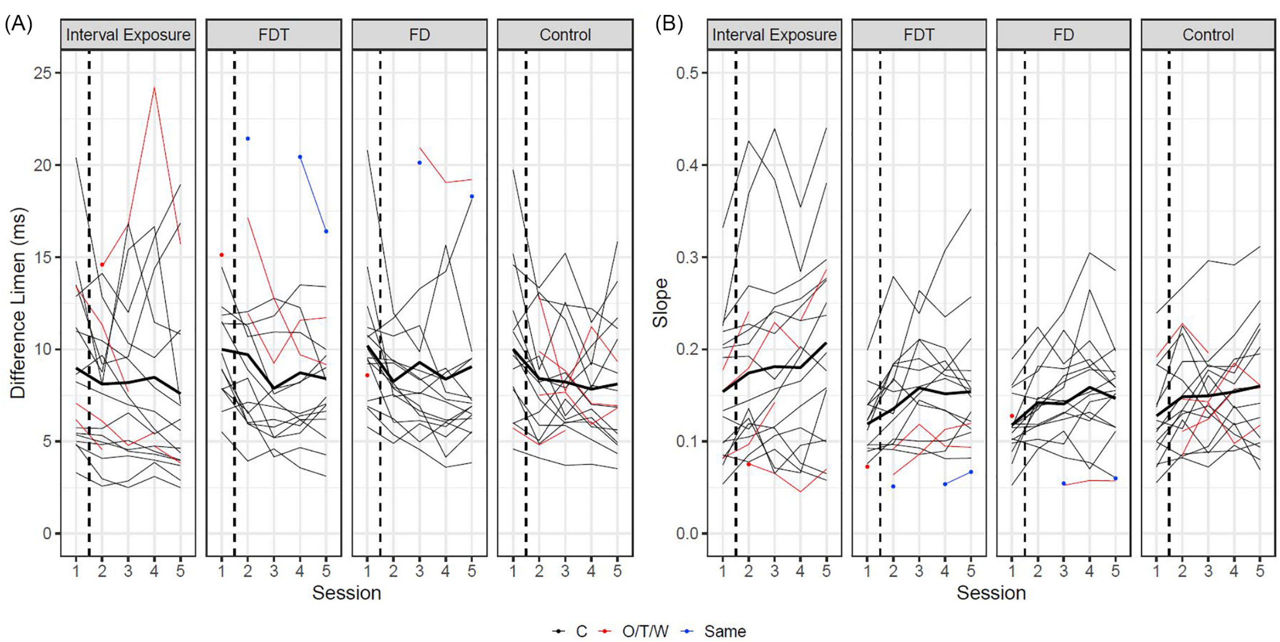

FIG. 5. (Color online) Subject level changes in (A) DLs and (B) slopes for temporal-interval discrimination across sessions. *C*, complete data; O/T/W, missing ≥1 session due to deviant performance, technical error, or withdrawal; same, same participant in group. Interval exposure = auditory temporal-interval discrimination (n_{total} = 20); FDT (n_{total} = 17); FD (n_{total} = 18); control = no exposure (n_{total} = 18). The bold solid lines represent group means. The vertical dashed line separates interval-training during exposure from interval target task training.

https://doi.org/10.1121/10.0017548

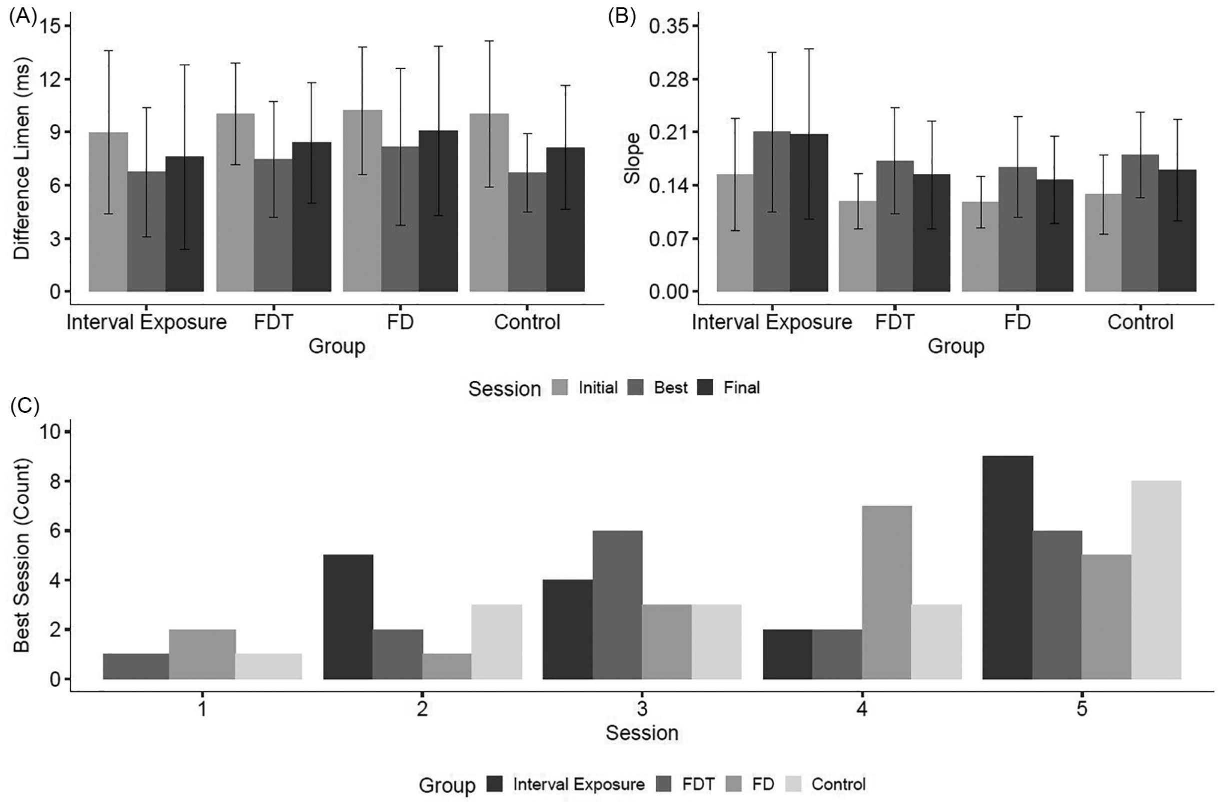

FIG. 6. (A) DLs and (B) slopes for the initial, best, and final sessions. Interval = auditory temporal-interval discrimination ($n_{\text{initial,best,final}}$ = 18,20,17); FDT ($n_{\text{initial,best,final}}$ = 14,16,16); FD ($n_{\text{initial,best,final}}$ = 16,17,17); control = no exposure ($n_{\text{initial,best,final}}$ = 15,18,17). The error bars represent SD. (C) The count of best performance at each session per group. This single plot represents best performance as defined by DLs and slopes, both of which produced identical counts.

$p = 0.934$, $\eta^2_p = 0.008$. Most groups exhibited best duration DLs at session 5. Half of the participants in the FDT group and most participants in the FD group exhibited best duration DLs at sessions 3 and 4, respectively.

Analyses of the slopes returned identical conclusions. A mixed ANOVA comparing the slopes from the initial to best sessions revealed a significant effect of session, $F(1,57) = 130.806$, $p = 2.18 \times 10^{-16}$, $\eta^2_p = 0.696$, but not group, $F(3,57) = 1.550$, $p = 0.212$, $\eta^2_p = 0.075$. The interaction was not significant, $F(3,57) = 0.346$, $p = 0.792$, $\eta^2_p = 0.018$. Comparisons of the slopes from the initial to final sessions revealed a significant effect of session, $F(1,54) = 44.090$, $p = 1.58 \times 10^{-8}$, $\eta^2_p = 0.449$, but not group, $F(3,54) = 2.042$, $p = 0.119$, $\eta^2_p = 0.102$, and interaction, $F(3,54) = 0.354$, $p = 0.786$, $\eta^2_p = 0.019$.

Figure 7 visualizes the change in performance on the target interval task. Difference scores between the initial and final sessions were calculated for duration DLs

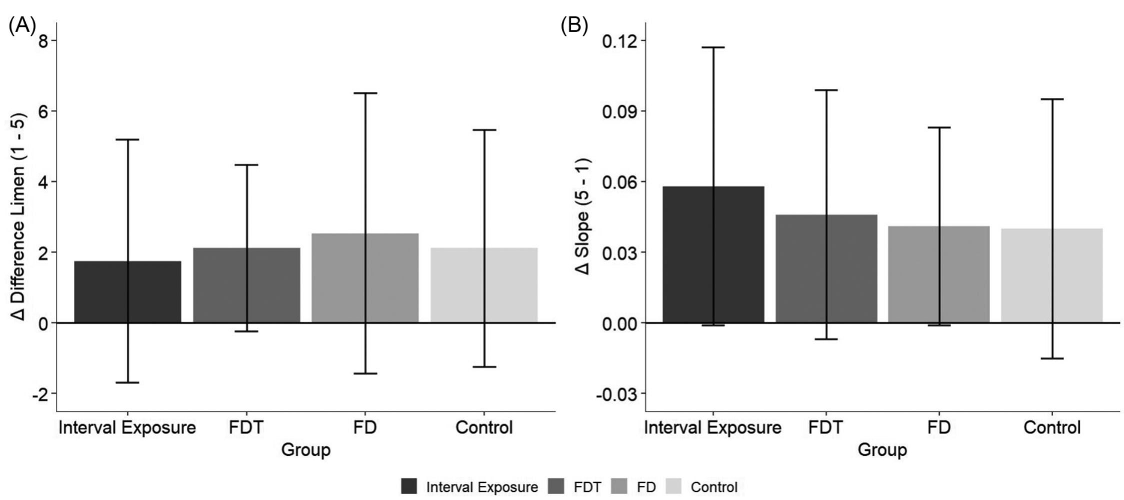

FIG. 7. Change in DLs (A) and slopes (B) at the initial and final sessions of the interval target task. Interval exposure = auditory temporal-interval discrimination ($n_{\text{initial,final}}$ = 18,17); FDT ($n_{\text{initial,final}}$ = 14,16); FD ($n_{\text{initial,final}}$ = 16,17); control = no exposure ($n_{\text{initial,final}}$ = 15,17). The error bars represent SD.

[Fig. 7(A)] and slopes [Fig. 7(B)]. Difference scores for duration DLs and slopes revealed that the interval and FDT groups improved the most across sessions, whereas the FD and control groups improved the least across sessions.

B. Effect of task-irrelevant temporal information on frequency discrimination

Although this study was not intended to evaluate frequency selectivity, we were interested in whether exposure to temporal-intervals during the FDT task enhanced frequency discrimination. A Welch two-samples *t*-test was used to compare frequency DLs (in Hz) between FDT-trained ($M = 15.804$, $SD = 7.565$) and FD-trained ($M = 18.787$, $SD = 10.441$) listeners. Although the FDT-trained listeners had lower frequency DLs than those for the FD-trained listeners, this difference was not statistically significant, $t(27.3) = 0.915$, $p = 0.368$, $\eta^2_p = 0.030$. Furthermore, the FDT-trained listeners ($M = 0.085$, $SD = 0.038$) had steeper slopes than those for the FD-trained listeners ($M = 0.079$, $SD = 0.046$), but this difference was not statistically significant, $t(28.7) = 0.442$, $p = 0.662$, $\eta^2_p = 0.007$.

V. DISCUSSION

Experience can enhance performance on tasks of temporal processing (Bratzke *et al.*, 2012; Bratzke *et al.*, 2014; Karmarkar and Buonomano, 2003; McGovern *et al.*, 2016; Wright *et al.*, 2010; Wright and Sabin, 2007; Xu *et al.*, 2021). Difficulty with isolating the individual effects of stimulus, procedure, and task experience has been a persistent limitation of the perceptual learning literature. Many studies of perceptual learning implement pre/posttest designs or some form of pretraining (Delhommeau *et al.*, 2002; Karmarkar and Buonomano, 2003; Lapid *et al.*, 2009; McGovern *et al.*, 2016) that could mask the contributions of procedural and task learning. The present study used an exposure design to investigate the role of perceptual, procedural, and task learning on temporal learning. Overall, temporal-interval discrimination improved, but an exposure (group) effect was not found when examined across sessions.

The findings of this study do not appear to be consistent with those from Ortiz and Wright (2009), who suggested that stimulus and procedure, but not task, experience contributes to overall learning. Several differences between the present study and that by Ortiz and Wright (2009) might explain the inconsistent findings. For one, the latter study's target task was an interaural timing difference task, which might involve different learning patterns than those needed for temporal-interval discrimination. Furthermore, their study estimated DLs using a two-alternative forced choice method, which has been found to produce DL estimates that are approximately 50% larger than those obtained from the method of constant stimuli applied in the present study (Ulrich, 2010). Inconsistencies in findings between the two studies might reflect differences in sensitivities to each of the two DL estimation methods. Last, their study design resulted in interpretations of perceptual, procedural, and task learning that differed from those in the present study. For example, during their exposure phase, a group of listeners was trained on the same task as their target task to impart information about the target task's stimulus, procedure, and task. The difference between this group's target task performance and that for another group that was exposed to only the target task's procedure and task was interpreted as perceptual learning. If there are unobserved factors beyond stimulus, procedure, and task that contributed to learning, the group exposed to the same task as the target task will likely perform best, which can lead to overestimating perceptual learning. The same limitation applies to the task learning in the present study, which was interpreted as the difference in improvement between the interval exposure- and FDT-trained listeners (see Table I). Future research is needed to identify these unobserved factors and the extent to which they contribute to temporal learning.

Also, the findings from the current study do not appear to be consistent with those from Xu *et al.* (2021), who found perceptual, procedural, and task contributions to overall learning on temporal-interval discrimination. These inconsistencies might stem from differences in study design. For example, their procedural learning was assessed through manipulating the statistical regularity of comparison intervals, whereas the present study manipulated listeners' experiences to the interval target task's procedures using an exposure phase. Further, the methods that Xu *et al.* (2021) used to estimate DLs varied from their trained bisection task (which used the method of constant stimuli) to their transfer comparison task (which used an adaptive staircase). Differences in sensitivities to each of the two DL estimation procedures could have biased their conclusions on task specificity to their trained bisection task. Last, the comparison task used in Xu *et al.* (2021) is more time-consuming to administer than their bisection task because it requires presenting the standard interval in every trial as opposed to only once at the start of each session. Consequently, participants might have been more fatigued during the comparison task than the bisection task, which could have masked possible effects of learning transfer and contributed to the observed task specificity.

Fatigue, boredom, and other factors might have contributed to the significant effect of time but not group in the current study. A recent study from Zhao *et al.* (2022) examined the role of fatigue, motivation, apathy, and confidence on several classic psychophysical tests calibrated using a stable and robust volume (i.e., level) setting procedure for online testing. They reported moderate, positive correlations between self-reported fatigue ratings and tone detection in noise. Self-ratings of confidence and metrics of apathy and motivation obtained using the Apathy Motivation Index questionnaire did not influence tone-in-noise thresholds. Further, high motivation, but not low fatigue, predicted an increased probe-signal effect (an accuracy advantage in detecting frequent 1000 Hz tones over infrequent, non-1000 Hz tones).

The present study was designed such that interval target task training immediately followed exposure to elements of the interval target task. Although the intent was to minimize the decay of acquired information between the exposure and training phases (Wright *et al.*, 2010), this required the three exposure groups to complete 1280 trials (640 from the exposure phase and 640 from the interval target task) in one sitting. Therefore, the three exposure groups could have become bored and inattentive and/or fatigued, especially after having to forgo in-laboratory administration for an entirely online format. These factors might have masked possible transfer of learning to the interval target task. This account is especially likely for the FDT and FD groups, which demonstrated best performance before the final session [see Fig. 6(C)]. An interesting line of future research could investigate the influences of information decay on learning transfer between exposure and interval target tasks and whether specific learning processes are more susceptible to this decay. This approach of research could contribute to study designs that minimize information decay between sessions while mitigating fatigue.

That listeners improved in temporal-interval discrimination across sessions suggests that some learning had occurred. A possible explanation for the lack of an exposure effect is that stimulus exposures encountered during the interval, FDT, and FD tasks simply had no true effect on target temporal-interval discrimination. This is unlikely because all exposure groups demonstrated larger numeric changes in slope from initial to final sessions than naive controls without any prior experience. Another possibility is that potential exposure effects were masked because participants were overtrained with 640 trials per session of temporal-interval discrimination. Wright and Sabin (2007) reported that training with 360 trials per day was sufficient to induce perceptual learning on auditory temporal-interval discrimination, and training with 900 trials per day did not yield additional learning. However, the two-alternative forced choice procedure used by Wright and Sabin (2007) has been found to produce larger DL estimates than those obtained using the method of constant stimuli, which employed in this study (Ulrich, 2010). Further, the adaptive staircase used in Wright and Sabin (2007) requires less trials to estimate DLs than the method of constant stimuli. Comparisons across the two studies are difficult given these differences in DL estimation. The temporal-interval discrimination task used in here was adapted from Bratzke *et al.* (2012), who applied nearly identical procedures to demonstrate substantial learning across sessions.

Several factors might substantiate the exposure-related enhancements to temporal-interval discrimination. For one, the interval exposure, FDT, FD, and interval target tasks shared identical procedures, including but not limited to the user interface, testing environment, the statistical regularity of the percepts of interest, and DL estimation procedure using the method of constant stimuli. For another, the FDT task was designed such that the low comparison frequency pairs (<1 kHz) always were copresented with short temporal-intervals (<100 ms; see Sec. III) and the high comparison frequency pairs (>1 kHz) always were copresented with long temporal-intervals (>100 ms). As such, temporal-intervals always were consistently coupled with the FDT task's frequency targets (low versus high). Despite not being statistically significant likely because of the large observed variances, frequency DLs for the FDT group were approximately 3 Hz lower (better) than those for the FD group, suggesting that at least some participants exploited the frequency-time relationship to benefit frequency discrimination performance. Successful recognition of the frequency-time associations might have promoted learning for the temporal-intervals, even though FDT-trained listeners were instructed to attend to frequency, not duration. This account is reminiscent of others used to explain learning of task-irrelevant information (Gabay *et al.*, 2015; Lim and Holt, 2011; Seitz and Watanabe, 2003, 2009; Vlahou *et al.*, 2012; Vlahou *et al.*, 2019; Wade and Holt, 2005; Wiener *et al.*, 2019), but the design of this study does not allow us to rule out the possibility that the additional passive exposures to temporal-intervals can enhance learning (Amitay *et al.*, 2006).

The entirely online format has several implications for the role of procedural learning that might explain the large observed variances (see Figs. 4–7). Despite that little is known about the extent to which online testing affects procedural learning for perceptual tasks, some predictions are possible. Compared to testing in controlled and presumably unfamiliar laboratory settings, completing experiments in familiar environments (e.g., the home) using familiar hardware (e.g., keyboard layout and computer monitor size) is expected to reduce the demand for procedural learning simply because there is less need to learn to interface with these factors. Concomitantly, completing online experiments imposes challenges on digital literacy and adaptability to new software. Further, noise levels (e.g., from sirens and vocalizations from other organisms) and the presence of other distractors (e.g., power outages) could fluctuate within and between sessions despite the measures taken to limit variability of the environment. These factors are expected to increase the demand for procedural learning.

A few listeners ($n = 6$) demonstrated worsening, defined as having >5 ms net increase in DL, in temporal-interval discrimination across sessions. This worsening effect has been observed in other studies involving temporal-interval discrimination (Wright *et al.*, 1997; Wright *et al.*, 2010; Xu *et al.*, 2021). In addition to boredom and fatigue, learning decay related to the time between sessions might have contributed to the worsening despite the measures taken to mitigate (see Sec. III). Still, the intent for using the multisession design was to maximize the potential for temporal perceptual learning and account for differences between the three learning processes later in time when perceptual learning is traditionally thought to dominate over procedural and task learning (Ortiz and Wright, 2010). Multisession designs provide numerous opportunities to retrieve acquired information for rehearsal and reconsolidation, which is thought to strengthen representations of the target precept through additional trials of training (Lee, 2008). This line of thought resonates with the literature, suggesting that perceptual learning can be enhanced after distributing and consolidating learning across time (Alain *et al.*, 2015; Savion-Lemieux and Penhune, 2005; Wright *et al.*, 1997). Specifically, distributing training-related exposure to stimulus features over several hundreds of trials across days is thought to facilitate neural plasticity in the sensory cortex, enlarging cortical representation for the trained percepts (Moore *et al.*, 2003; Recanzone *et al.*, 1992; Recanzone *et al.*, 1993. Sleep, which presumably occurred between sessions, is thought to enhance this auditory learning consolidation (Atienza *et al.*, 2004; Chen *et al.*, 2017; Gaab *et al.*, 2004; Gottselig *et al.*, 2004).

Overall, stimulus, procedure, and task contributions to temporal-interval discrimination appear to be negligible. The large observed variances likely masked stimulus, procedure, and task contributions toward overall temporal learning. Consequently, this study cannot rule out the notion that multiple distinct processes are engaged to varying extents during learning (Ahissar and Hochstein, 2004; Cohen *et al.*, 2013; Ortiz and Wright, 2009; Xu *et al.*, 2021). Caution is needed when designing studies of temporal perceptual learning to avoid biasing conclusions about changes to the perceptual system related to feature training alone. Finally, given the null results and large observed variances (see Fig. 4–7), this study does not discount contributions from factors beyond stimulus, procedure, and task on temporal learning.

ACKNOWLEDGMENTS

We thank Dr. David McGovern for sharing the MATLAB scripts from his experiment and Waris Aiemworawutikul for providing feedback on programming with PsychoPy and setting up alternative web hosting when the website[4] became unavailable because of a server fire. The project described was supported by the National Institutes of Health through Grant Nos. UL1 TR001857, KL2 TR001856, and/or TL1 TR001858.

[1]See https://hearingtest.online/ (Last viewed May 8, 2021).
[2]See https://www.essilor.com/en/vision-tests/test-your-vision/ (Last viewed May 8, 2021).
[3]See https://github.com/lzhen4/mle (Last viewed November 26, 2022).
[4]See https://pavlovia.org (Last viewed May 15, 2021).

Ahissar, M., and Hochstein, S. (**1993**). "Attentional control of early perceptual learning," Proc. Natl. Acad. Sci. U.S.A. **90**(12), 5718–5722.
Ahissar, M., and Hochstein, S. (**1997**). "Task difficulty and the specificity of perceptual learning," Nature **387**(6631), 401–406.
Ahissar, M., and Hochstein, S. (**2004**). "The reverse hierarchy theory of visual perceptual learning," Trends Cognit. Sci. **8**(10), 457–464.
Ahissar, M., Nahum, M., Nelken, I., and Hochstein, S. (**2009**). "Reverse hierarchies and sensory learning," Philos. Trans. R. Soc. B **364**(1515), 285–299.
Alain, C., Zhu, K. D., He, Y., and Ross, B. (**2015**). "Sleep-dependent neuroplastic changes during auditory perceptual learning," Neurobiol. Learn. Mem. **118**, 133–142.
American Speech-Language-Hearing Association (**2005**). "Guidelines for manual pure-tone threshold audiometry [Guidelines]," available at https://www.asha.org/policy/GL2005-00014/ (Last viewed September 20, 2019).
Amitay, S., Irwin, A., and Moore, D. R. (**2006**). "Discrimination learning induced by training with identical stimuli," Nat. Neurosci. **9**(11), 1446–1448.
Amitay, S., Zhang, Y. X., and Moore, D. R. (**2012**). "Asymmetric transfer of auditory perceptual learning," Front. Psychol. **3**, 508.
Atienza, M., Cantero, J. L., and Stickgold, R. (**2004**). "Posttraining sleep enhances automaticity in perceptual discrimination," J. Cogn. Neurosci. **16**(1), 53–64.
Bratzke, D., Schröter, H., and Ulrich, R. (**2014**). "The role of consolidation for perceptual learning in temporal discrimination within and across modalities," Acta Psychol. **147**, 75–79.
Bratzke, D., Seifried, T., and Ulrich, R. (**2012**). "Perceptual learning in temporal discrimination: Asymmetric cross-modal transfer from audition to vision," Exp. Brain Res. **221**(2), 205–210.
Bruns, P., and Watanabe, T. (**2019**). "Perceptual learning of task-irrelevant features depends on the sensory context," Sci. Rep. **9**(1), 1666.
Chen, L., Guo, L., and Bao, M. (**2017**). "Sleep-dependent consolidation benefits fast transfer of time interval training," Exp. Brain Res. **235**(3), 661–672.
Cohen, Y., Daikhin, L., and Ahissar, M. (**2013**). "Perceptual learning is specific to the trained structure of information," J. Cogn. Neurosci. **25**(12), 2047–2060.
Delhommeau, K., Micheyl, C., Jouvent, R., and Collet, L. (**2002**). "Transfer of learning across durations and ears in auditory frequency discrimination," Percept. Psychophys. **64**(3), 426–436.
Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. (**2007**). "G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences," Behav. Res. Methods **39**(2), 175–191.
Gaab, N., Paetzold, M., Becker, M., Walker, M. P., and Schlaug, G. (**2004**). "The influence of sleep on auditory learning: A behavioral study," Neuroreport **15**(4), 731–734.
Gabay, Y., Dick, F. K., Zevin, J. D., and Holt, L. L. (**2015**). "Incidental auditory category learning," J. Exp. Psychol.: Hum. Percept. Perform. **41**(4), 1124–1138.
Gottselig, J. M., Hofer-Tinguely, G., Borbély, A. A., Regel, S. J., Landolt, H. P., Rétey, J. V., and Achermann, P. (**2004**). "Sleep and rest facilitate auditory learning," Neurosci. **127**(3), 557–561.
Hawkey, D. J., Amitay, S., and Moore, D. R. (**2004**). "Early and rapid perceptual learning," Nat. Neurosci. **7**(10), 1055–1056.
Hochstein, S., and Ahissar, M. (**2002**). "View from the top: Hierarchies and reverse hierarchies in the visual system," Neuron **36**(5), 791–804.
Hussain, Z., McGraw, P. V., Sekuler, A. B., and Bennett, P. J. (**2012**). "The rapid emergence of stimulus specific perceptual learning," Front. Psychol. **3**, 226.
Jeter, P. E., Dosher, B. A., Petrov, A., and Lu, Z. L. (**2009**). "Task precision at transfer determines specificity of perceptual learning," J. Vision **9**(3), 1–13.
Karmarkar, U. R., and Buonomano, D. V. (**2003**). "Temporal specificity of perceptual learning in an auditory discrimination task," Learn. Mem. **10**(2), 141–147.
Lapid, E., Ulrich, R., and Rammsayer, T. (**2009**). "Perceptual learning in auditory temporal discrimination: No evidence for a cross-modal transfer to the visual modality," Psychon. Bull. Rev. **16**(2), 382–389.
Leclercq, V., and Seitz, A. R. (**2012**). "The impact of orienting attention in fast task-irrelevant perceptual learning," Atten. Percept. Psychophys. **74**(4), 648–660.
Lee, J. L. (**2008**). "Memory reconsolidation mediates the strengthening of memories by additional learning," Nat. Neurosci. **11**(11), 1264–1266.
Lim, S. J., and Holt, L. L. (**2011**). "Learning foreign sounds in an alien world: Videogame training improves non-native speech categorization," Cognit. Sci. **35**(7), 1390–1405.
Linares, D., and López-Moliner, J. (**2016**). "quickpsy: An *R* package to fit psychometric functions for multiple groups," R J. **8**(1), 122–131.
Liu, Z., and Weinshall, D. (**2000**). "Mechanisms of generalization in perceptual learning," Vision Res. **40**(1), 97–109.
McGovern, D. P., Astle, A. T., Clavin, S. L., and Newell, F. N. (**2016**). "Task-specific transfer of perceptual learning across sensory modalities," Curr. Biol. **26**(1), R20–R21.
Meuwese, J. D., Post, R. A., Scholte, H. S., and Lamme, V. A. (**2013**). "Does perceptual learning require consciousness or attention?," J. Cogn. Neurosci. **25**(10), 1579–1596.
Moore, D. R., Amitay, S., and Hawkey, D. J. (**2003**). "Auditory perceptual learning," Learn. Mem. **10**(2), 83–85.
NIOSH (**2020**). "NIOSH Sound Level Meter App (1.2.4) [Mobile App]," National Institute for Occupational Safety and Health, available at https://apps.apple.com/us/app/niosh-slm/id1096545820 (Last viewed May 14, 2021).

Ortiz, J. A., and Wright, B. A. (**2009**). "Contributions of procedure and stimulus learning to early, rapid perceptual improvements," J. Exp. Psychol.: Hum. Percept. Perform. **35**(1), 188–194.
Ortiz, J. A., and Wright, B. A. (**2010**). "Differential rates of consolidation of conceptual and stimulus learning following training on an auditory skill," Exp. Brain Res. **201**(3), 441–451.
Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., and Lindeløv, J. K. (**2019**). "PsychoPy2: Experiments in behavior made easy," Behav. Res. **51**(1), 195–203.
Recanzone, G. H., Merzenich, M. M., Jenkins, W. M., Grajski, K. A., and Dinse, H. R. (**1992**). "Topographic reorganization of the hand representation in cortical area 3b owl monkeys trained in a frequency-discrimination task," J. Neurophysiol. **67**(5), 1031–1056.
Recanzone, G. H., Schreiner, C. E., and Merzenich, M. M. (**1993**). "Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys," J. Neurosci. **13**(1), 87–103.
Roelfsema, P. R., van Ooyen, A., and Watanabe, T. (**2010**). "Perceptual learning rules based on reinforcers and attention," Trends Cognit. Sci. **14**(2), 64–71.
Savion-Lemieux, T., and Penhune, V. B. (**2005**). "The effects of practice and delay on motor skill learning and retention," Exp. Brain Res. **161**(4), 423–431.
Seitz, A., Lefebvre, C., Watanabe, T., and Jolicoeur, P. (**2005**). "Requirement for high-level processing in subliminal learning," Current Biol. **15**(18), R753–R755.
Seitz, A., and Watanabe, T. (**2005**). "A unified model for perceptual learning," Trends Cognit. Sci. **9**(7), 329–334.
Seitz, A. R., and Dinse, H. R. (**2007**). "A common framework for perceptual learning," Curr. Opin. Neurobiol. **17**(2), 148–153.
Seitz, A. R., Kim, D., and Watanabe, T. (**2009**). "Rewards evoke learning of unconsciously processed visual stimuli in adult humans," Neuron **61**(5), 700–707.
Seitz, A. R., Protopapas, A., Tsushima, Y., Vlahou, E. L., Gori, S., Grossberg, S., and Watanabe, T. (**2010**). "Unattended exposure to components of speech sounds yields same benefits as explicit auditory training," Cognition **115**(3), 435–443.
Seitz, A. R., and Watanabe, T. (**2003**). "Psychophysics: Is subliminal learning really passive?," Nature **422**(6927), 36.
Seitz, A. R., and Watanabe, T. (**2008**). "Is task-irrelevant learning really task-irrelevant?," PloS One **3**(11), e3792.
Seitz, A. R., and Watanabe, T. (**2009**). "The phenomenon of task-irrelevant perceptual learning," Vision Res. **49**(21), 2604–2610.
SkyPaw Co., Ltd. (**2021**). Decibel X–dB Sound Level Meter, Noise Detector (6.3.4) [Mobile App], available at https://play.google.com/store/apps/details?id=com.skypaw.decibel&hl=en&gl=US (Last viewed May 15, 2021).
Szpiro, S. F., Wright, B. A., and Carrasco, M. (**2014**). "Learning one task by interleaving practice with another task," Vision Res. **101**, 118–124.
Ulrich, R. (**2010**). "DLs in reminder and 2AFC tasks: Data and models," Atten. Percept. Psychophys. **72**(4), 1179–1198.
Vlahou, E., Seitz, A. R., and Kopčo, N. (**2019**). "Nonnative implicit phonetic training in multiple reverberant environments," Atten. Percept. Psychophys. **81**(4), 935–947.
Vlahou, E. L., Protopapas, A., and Seitz, A. R. (**2012**). "Implicit training of nonnative speech stimuli," J. Exp. Psychol. General **141**(2), 363–381.
Wade, T., and Holt, L. L. (**2005**). "Incidental categorization of spectrally complex non-invariant auditory stimuli in a computer game task," J. Acoust. Soc. Am. **118**(4), 2618–2633.
Watanabe, T., Náñez, J. E., and Sasaki, Y. (**2001**). "Perceptual learning without perception," Nature **413**(6858), 844–848.
Watanabe, T., and Sasaki, Y. (**2015**). "Perceptual learning: Toward a comprehensive theory," Annu. Rev. Psychol. **66**, 197–221.
Wiener, S., Murphy, T. K., Goel, A., Christel, M. G., and Holt, L. L. (**2019**). "Incidental learning of non-speech auditory analogs scaffolds second language learners' perception and production of Mandarin lexical tones," in *Proceedings of the International Conference on Phonetic Sciences*, pp. 1699–1703.
Wiggs, C. L., and Martin, A. (**1998**). "Properties and mechanisms of perceptual priming," Curr. Opin. Neurobiol. **8**(2), 227–233.
Wright, B. A., Buonomano, D. V., Mahncke, H. W., and Merzenich, M. M. (**1997**). "Learning and generalization of auditory temporal-interval discrimination in humans," J. Neurosci. **17**(10), 3956–3963.
Wright, B. A., and Fitzgerald, M. B. (**2001**). "Different patterns of human discrimination learning for two interaural cues to sound-source location," Proc. Natl. Acad. Sci. U.S.A. **98**(21), 12307–12312.
Wright, B. A., and Sabin, A. T. (**2007**). "Perceptual learning: How much daily training is enough?," Exp. Brain Res. **180**(4), 727–736.
Wright, B. A., Sabin, A. T., Zhang, Y., Marrone, N., and Fitzgerald, M. B. (**2010**). "Enhancing perceptual learning by combining practice with periods of additional sensory stimulation," J. Neurosci. **30**(38), 12868–12877.
Wright, B. A., and Zhang, Y. (**2009**). "A review of the generalization of auditory learning," Philos. Trans. R. Soc. B **364**(1515), 301–311.
Xu, R., Church, R. M., Sasaki, Y., and Watanabe, T. (**2021**). "Effects of stimulus and task structure on temporal perceptual learning," Sci. Rep. **11**(1), 668.
Zhao, S., Brown, C. A., Holt, L. L., and Dick, F. (**2022**). "Robust and efficient online auditory psychophysics," Trends Hear. **26**, 1–24.