Attention, Perception, & Psychophysics
https://doi.org/10.3758/s13414-022-02616-x

Incidental auditory category learning and visuomotor sequence learning do not compete for cognitive resources

Yafit Gabay[1,2] · Michelle Madlansacay[3] · Lori L. Holt[3,4]

Accepted: 3 November 2022

Abstract

The environment provides multiple regularities that might be useful in guiding behavior if one was able to learn their structure. Understanding statistical learning across simultaneous regularities is important, but poorly understood. We investigate learning across two domains: visuomotor sequence learning through the serial reaction time (SRT) task, and incidental auditory category learning via the systematic multimodal association reaction time (SMART) task. Several commonalities raise the possibility that these two learning phenomena may draw on common cognitive resources and neural networks. In each, participants are uninformed of the regularities that they come to use to guide actions, the outcomes of which may provide a form of *internal feedback*. We used dual-task conditions to compare learning of the regularities in isolation versus when they are simultaneously available to support behavior on a seemingly orthogonal visuomotor task. Learning occurred across the simultaneous regularities, without attenuation even when the informational value of a regularity was reduced by the presence of the additional, convergent regularity. Thus, the simultaneous regularities do not compete for associative strength, as in overshadowing effects. Moreover, the visuomotor sequence learning and incidental auditory category learning do not appear to compete for common cognitive resources; learning across the simultaneous regularities was comparable to learning each regularity in isolation.

Keywords Incidental auditory category learning · Dual task · Visuomotor sequence learning · Overshadowing · Statistical learning

Introduction

The natural world presents regularities that might guide successful behavior if we can learn their structure. Quite often, this learning proceeds without explicit instruction or feedback. Moreover, multiple – often simultaneous – input dimensions frequently convey information that might collaborate to support behavior. As an example, complex acoustic information

Significance statement The environment presents multiple regularities that might be useful in guiding behavior if we can learn their structure. Simultaneous regularities may facilitate learning, or interfere with learning if learning in each domain competes for limited, common cognitive resources, or if the combination of regularities lowers the informational value with which each regularity predicts behavior. In this way, comparison of learning across simultaneous regularities versus learning of the regularities in isolation informs our understanding of whether common processes are involved across learning challenges. Here, we observe that incidental learning of auditory categories and visuomotor sequence learning – two tasks with multiple commonalities – do not interfere with one another when both regularities exist to inform behavior.

✉ Yafit Gabay
ygabay@edu.haifa.ac.il

1 Department of Special Education, University of Haifa, Haifa, Israel

2 Edmond J. Safra Brain Research Center for the Study of Learning Disabilities, 199 Abba Khoushy Ave, Haifa, Israel

3 Department of Psychology, Carnegie Mellon University, Pittsburgh, PA, USA

4 Neuroscience Institute, Carnegie Mellon University, Pittsburgh, PA, USA

Published online: 12 December 2022

that conveys whether an infant's cry is hunger or pain aligns with a sequence of the infant's facial expressions. Each information source is variable, yet each provides structured regularity to direct a caregiver's behavior and to hone that behavior with experience.

Understanding how learning proceeds across simultaneous input regularities can shed light on how observers learn input regularities in complex environments and can help us to differentiate among contributing mechanisms. Multiple regularities may align to provide joint information that facilitates learning. Or, simultaneous regularities may interfere with learning if learning in each domain competes for limited, common cognitive resources, or if the combination of regularities lowers the informational value with which each regularity predicts behavior. There are many demonstrations that adult learners can track multiple statistical regularities simultaneously and use them to guide behavior (Conway & Christiansen, 2006; Conway et al., 2020; Deocampo et al., 2019; Goschke & Bolte, 2012; Iao et al., 2021; Idemaru & Holt, 2014; Jiménez & Vázquez, 2011; Mitchel & Weiss, 2011; Vuong et al., 2016; Zhang & Holt, 2018). However, less is known about the extent to which learning across multiple regularities may introduce interference, for example in competition for common cognitive resources.

Here, we draw from two distinct learning paradigms, incidental auditory category learning and visuomotor sequence learning, to investigate the information value of simultaneous regularities that predict a common behavioral outcome. As we describe below, these two learning domains share multiple commonalities that, at least on a surface level, suggest the possibility that they might engage common learning mechanisms, or draw upon common cognitive resources. However, these possibilities have not been explicitly examined.

Here, we characterize learning of each regularity in isolation, and also when the two align to guide behavior. On the one hand, simultaneous, redundant regularities may facilitate learning above and beyond learning each regularity in isolation. On the other hand, the availability of multiple regularities creates circumstances in which they may interfere with learning, for example, if the learning draws upon common cognitive resources or when the combination of regularities lowers the informational value with which each predicts behavior.

With regard to this latter point, learners are sensitive to correlations among events and are affected by informational value – the capacity of a stimulus to act as a reliable predictor of an outcome relative to other stimuli (McLaren & Dickinson, 1990). This is best demonstrated in cue competition effects such as blocking and overshadowing, in which cues compete for associative strength (Mackintosh, 1975). For example, if two cues (X, Y) jointly predict an outcome the subsequent response to either X or Y is smaller than the response had X and Y been learned in isolation (Pavlov, 1927). Associative learning models reliant on a single prediction-error term (Rescorla, 1972) predict that the associative strength of a regularity will be reduced in the context of learning simultaneous regularities compared to learning the regularity in isolation. This phenomenon is described as overshadowing and has been observed across species (Kamin, 1968; Tennant & Bitterman, 1975) as well as across a variety of tasks and domains (Chamizo et al., 2003; Chapman & Robbins, 1990; Ellis, 2006; Gluck & Bower, 1988; Hou, 2021) including category learning (Lau et al., 2020; Nixon, 2020). Yet, the extent to which this effect exists in human *incidental* learning – of the type involved in visuomotor sequence learning and incidental auditory category learning examined here – is still a matter of debate (Beesley & Shanks, 2012; McLaren et al., 2014; Schmidt & De Houwer, 2019).

Incidental auditory category learning Learning to treat statistically structured distributions of distinct objects or events as functionally equivalent members of a category is crucial to effective behavior. In natural environments, category learning tends to proceed across multiple, simultaneously present forms of input typically with no explicit instruction or overt feedback as a guide. Prior research demonstrates that speech and other auditory categories can be learned *incidentally* across statistically structured input even when individuals are unaware that categories exist, do not make overt category decisions, and do not receive explicit feedback (Gabay et al., 2015; Lim et al., 2019; Lim et al., 2015; Lim & Holt, 2011; Roark et al., 2021; Seitz et al., 2010; Vlahou et al., 2012; Wade & Holt, 2005). This learning occurs when statistically structured input exemplars align with behaviorally relevant events and responses in a primary task ostensibly unrelated to category learning (Gabay et al., 2015; Roark et al., 2021).

For example, in the Systematic Multi-Modal Association Response Time (SMART) task, participants rapidly report the location of a visual target in one of four possible screen locations (Gabay et al., 2015). Acoustically variable sound exemplars from one of four novel categories precede the appearance of a visual target and each auditory category aligns with one visual target location, predicting where the target will appear only if one learns to categorize the variable sounds. Participants are not informed about the alignment of auditory categories to visual target location and can perform the visuomotor task without sound. Nonetheless, the relationship between the complex acoustic regularity defining the sound categories and the upcoming location of the visual target results in auditory category learning that generalizes robustly to novel exemplars (Gabay et al., 2015; Gabay et al., 2019; Roark et al., 2021; Wade & Holt, 2005). Accordingly, elimination of the category-to-location alignment produces a significant slowing in the time to detect the visual target – a response-time (RT) Cost – indicative of a reliance on auditory categorization to facilitate speedy response to the visual target. This is not a simple auditory-to-visual association. The mapping from auditory categories to visual locations is many-to-one, and incidental category learning generalizes to overt labeling of novel category exemplars not experienced in the SMART task. Adult participants acquire novel auditory categories via incidental learning (Gabay et al., 2015; Wade & Holt, 2005) that are not learned across passive exposure or unsupervised sorting (Emberson et al., 2013; Roark et al., 2021; Wade & Holt, 2005); this learning is dependent upon categories being defined by acoustic regularities, even if complex (Gabay et al., 2015; Lim, Fiez & Holt, 2019), and by alignment of these regularities with a unique motor response (Roark et al., 2021). At the neural level, incidental auditory category learning has been associated with activation of cortico-striatal loops (posterior striatum, in particular the body of the caudate and putamen, with left posterior superior temporal sulcus; Lim, Fiez, & Holt, 2019).

Visuomotor sequence learning The SMART paradigm shares commonalities with the Serial Reaction Time Task (SRT), one of the most widely used paradigms for examining visuomotor sequence learning (Abrahamse et al., 2010; Cleeremans, 1997; Janacsek et al., 2012; Nissen & Bullemer, 1987; Schwarb & Schumacher, 2012; Szegedi-Hallgató et al., 2019). Like the SMART task, the SRT involves eliciting an action in response to a specific cue (e.g., a red X) and participants are unaware of the presence of an aligned statistical structure. In the SRT (but not in SMART), visual target locations follow a repeated pattern unknown to participants, such that the position of the next stimulus can be predicted from the previous one (Dezfouli & Balleine, 2012). Participants learn these sequences, as indicated by speedier RTs to report the position of the visual target for repeated sequences and longer RTs when a random sequence is introduced. After practice, participants behave in a predictive manner even in the absence of a visual cue (Matsuzaka et al., 2007). At the neural level, visuomotor sequence learning acquired via the SRT task has been associated with the visuomotor (the anterior putamen) and executive cortico-striatal loops (head of the caudate) (Hazeltine & Ivry, 2003; Janacsek et al., 2020; Peigneux et al., 2000).

Experiment 1

Several commonalities raise the possibility that both incidental auditory category learning and visuomotor sequence learning may draw on common cognitive resources. In each, participants are uninformed of the input regularities. Yet, the regularities come to support success in the primary task, guiding speedy action. The outcomes of these actions may provide a form of *internal feedback* derived from successfully using statistically structured input to guide predictions about upcoming actions. This differs from other forms of statistical learning whereby regularities are acquired through passive exposure (Saffran et al., 1996; Turk-Browne et al., 2005). Yet, despite surface commonalities across learning in the SMART and SRT paradigms, we do not know if they draw upon common resources.

There is a priori reason to posit that these two forms of learning may draw upon different processes. Incidental category learning via the SMART paradigm involves the acquisition of distributional statistics whereas motor sequence learning via the SRT task relates to the formation of conditional statistics, which have been proposed to rely upon distinct learning mechanisms (Thiessen et al., 2013). Furthermore, the SMART task involves learning across audio-visuomotor input, whereas learning in the SRT task establishes visuomotor associations; learning processes may play out differently across modalities (Conway, 2020; Frost et al., 2015). Neurobiologically, the striatum has been implicated in both learning challenges (Lim et al., 2019; Peigneux et al., 2000; Rauch et al., 1997). However, this broad characterization of a complex network is likely to mask important differentiation of striatal processing across tasks.

Here, we exploit the commonalities of these two learning challenges as a conservative test case in which redundant regularities present under highly similar task demands might be expected to draw upon common cognitive resources (Kahneman, 1973; Posner & Petersen, 1990; Tombu & Jolicœur, 2003) or to lower the informational value with which each predicts behavior and interferes with learning. Alternatively, distinct processes may support the ability to learn the dual regularities simultaneously, without diminishment relative to learning each regularity in isolation. Finally, there is also the possibility of mutual facilitation in learning the simultaneous regularities, compared to learning each regularity in isolation. For example, accumulating motor sequential knowledge may support predictions that scaffold auditory category learning. Here, we investigate these possibilities by examining incidental auditory category learning in SMART and visuomotor learning in SRT in isolation, and in combination.

Methods

Participants Young adult participants (104 total; 41 males and 61 females, M_{age} = 25 years, *SD* = 3.5 years; due to experimenter error demographics are not available for two participants) were recruited in-person and assigned randomly to one of four conditions: (1) *Category, No Sequence* (N = 25); (2) *Category + Sequence (Category Violation)* (N = 26); (3) *Sequence, No Category* (N = 25), (4) *Sequence + Category (Sequence Violation)* (N = 28). A power analysis (G*Power and the R *pwr* package; Faul et al. 2007) using the effect sizes from Experiments 1 and 3 of Gabay et al. (2015) revealed that a sample of 21 participants/group would establish statistical power at a .90 level with alpha of .05 to detect a difference across conditions (d = .87 or f = .4). The study was conducted at the University of Haifa in accordance with the tenets of the Declaration of Helsinki.

Stimuli Figure 1a illustrates four novel nonspeech auditory categories developed by Wade and Holt (2005) and used in prior studies (Gabay et al., 2015; Gabay & Holt, 2015; Leech et al., 2009; Liu & Holt, 2011; Roark et al., 2021; Wade & Holt, 2005). These sounds have some of the spectrotemporal complexity of speech, but are unequivocally nonspeech owing to their noise and square wave sources. Each category has six exemplars and five novel stimuli withheld from training to test generalization of category learning (not illustrated). A simple unidimensional acoustic cue defines two of the categories (the higher-frequency component increases or decreases in frequency; Fig. 1a, categories A and B). No one acoustic cue uniquely defines category membership in the other two, multidimensional, auditory categories, although exemplars exhibit regularity in higher-dimensional acoustic space (Fig. 1a, categories C and D; see Lim, Fiez, & Holt, 2019; Wade & Holt, 2005).

SMART task In the SMART task participants rapidly report the location of a visual target that appears in one of four possible

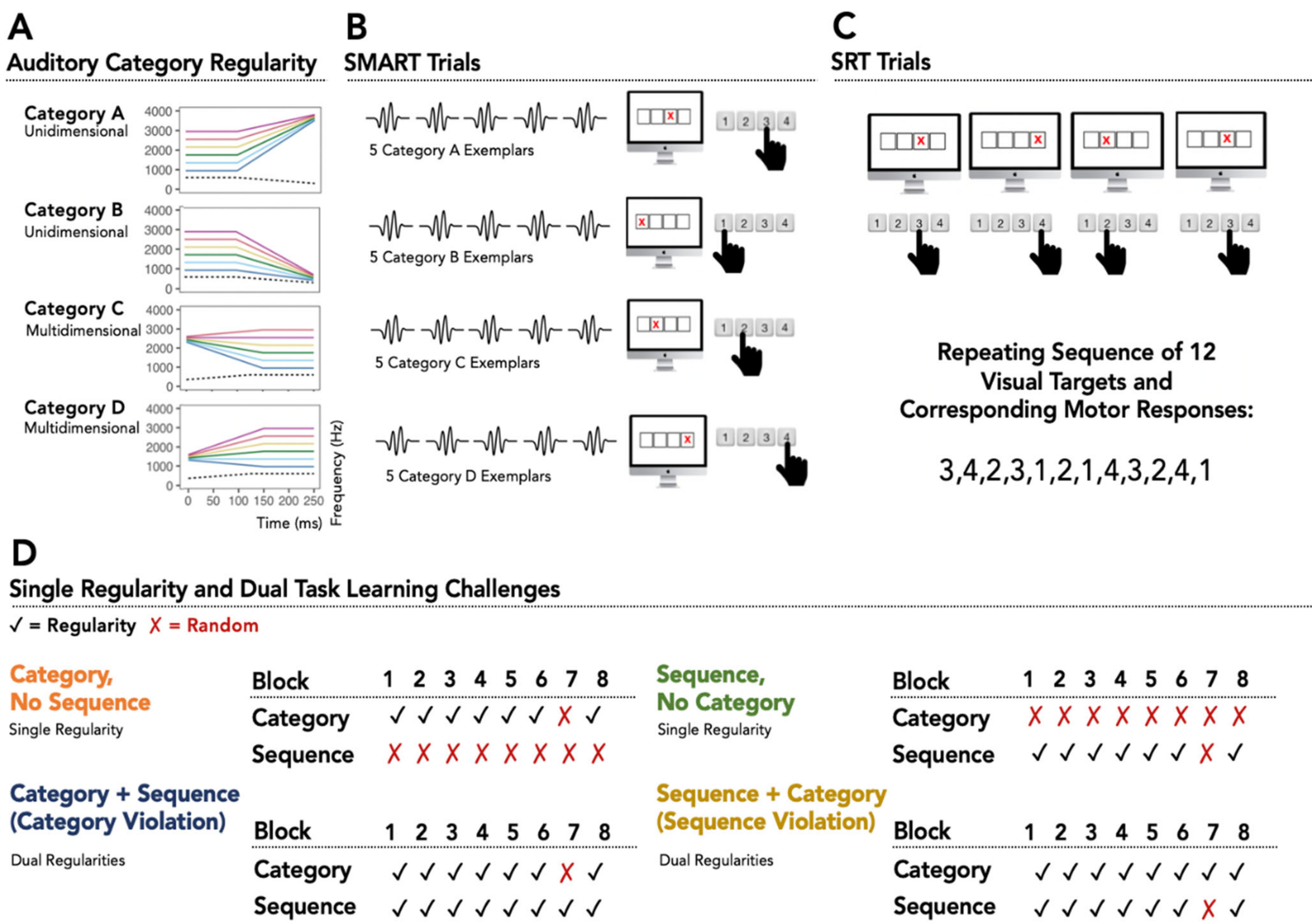

Fig. 1 Overview of stimuli and paradigm. **a** Four nonspeech auditory categories are defined across six exemplars (differentiated by the higher-frequency component shown as different colors on the same axes). Categories A and B are characterized by a unidimensional acoustic attribute (offset rises or falls), whereas Categories C and D cannot be defined by a single acoustic attribute and are defined across multiple acoustic dimensions in higher-order space (see Wade & Holt, 2005). **b** In the Systematic Multimodal Association Reaction Time (SMART) task each auditory category uniquely predicts the upcoming location of a visual target that is responded to with a unique button press to indicate target location. Here, the visual targets appear in a random order across trials. **c** In the Serial Reaction Time (SRT) task, the order of visual target follows a 12-item sequence. Participants press a unique key to indicate the location of the visual target. To parallel the SMART, the SRT task in this study included five acoustic exemplars preceding the visual target; in the Sequence, No Category condition there was no association of auditory category to visual target location. **d** Four conditions involve manipulation of the presence (black check) or absence (red X) of auditory category and visuomotor sequence regularities across eight blocks. Two conditions provide a single regularity for category (orange) or sequence (green). The other two conditions involve dual-task scenarios in which both category and sequence regularities are present, except in the block that tests for response time (RT) costs when a regularity is violated for category (blue) or sequence (yellow)

screen locations with a key press corresponding to location (Fig. 1b). A brief sequence of five sounds precedes each visual target. On each trial, five of six unique exemplars drawn from one of four auditory categories are randomly selected (without replacement) and presented in a random order with a 50-ms silent interval between sounds. This creates a many-to-one mapping such that multiple acoustically variable sound category exemplars are associated with a single visual location, both within and across trials. Participants are not informed about the relationship, and it is not necessary to successfully report visual target location. However, since sound categories perfectly predict upcoming visual target location *and* the corresponding response button to be pressed (Roark et al., 2021), learning to categorize acoustically variable sounds in predicting the location of an upcoming visual target can facilitate motor response on the primary task, without requiring overt sound categorization decisions or even awareness of the existence of auditory categories.

Monitoring RT to report the visual target location (from visual target onset) provides a covert measure of category learning. In the training blocks (Blocks 1–6 and 8), there is a perfect association of each sound category to a particular visual location. A test block scrambles the category-to-location mapping (Block 7). In Block 7, each trial is composed of five exemplars randomly selected from the pool of all exemplars, across categories. Thus, sounds are familiar but there is no category-to-location mapping and no within-trial category exemplar similarity. If participants learn the sound categories across training blocks and come to rely upon the category-to-location mapping to direct visuomotor behavior, then response to the visual target is expected be slower in the test block relative to the training block that preceded it (RT_{Block7}-RT_{Block6} = RT Cost).

A four-alternative, forced-choice post-test follows the SMART task. Here, all sounds are novel category exemplars not experienced in the SMART task so success requires generalization of incidental category learning. At post-test, a single novel sound exemplar repeats five times and participants guess the location where the visual target would be most likely to appear. No visual targets appear; there is no feedback. Accuracy, the proportion of trials (96 total trials drawn from five novel exemplars from each of four categories) for which the response matches the category-to-location mapping experienced in the SMART task, is expected to be above chance (> .25) if learning generalizes to novel exemplars, a hallmark of robust category learning.

SRT task Like the SMART task, the SRT task involves rapidly reporting the location of a visual target that appears in one of four possible screen locations with a key press corresponding to location. However, unknown to the participants, the visual target position appears in a repeating 12-trial complex second-order conditional sequence in which every location was determined by the previous two locations (342312143241). The sequence was balanced for location frequency (each location occurred three times), transition frequency (each possible transition from one location to a different one occurred once), and repetitions (no repetitions), as in previous studies (Gabay et al., 2012a, 2012b; Reed & Johnson, 1994). The sequence was repeated four times in each training block. In a test block (Block 7), the sequence was randomized. If participants learn the visuomotor sequence across training blocks and come to rely upon it to direct visuomotor behavior, then response to the visual target is expected to be slower in the test block relative to the training block that preceded it (RT_{Block7}-RT_{Block6} = RT Cost).

Single regularity and dual task learning challenges Two conditions examined incidental auditory category learning in the presence or absence of visuomotor sequence learning, as illustrated in Fig. 1c. The *Category, No Sequence* condition involved a consistent mapping between auditory category and visual target location, but no repeated visuomotor sequence across six training blocks, followed by a randomization of this mapping in Block 7 and a return to the category-to-location regularity in Block 8. A *Category + Sequence (Category Violation)* condition was identical to the *Category, No Sequence* condition, except that a 12-trial visuomotor sequence aligned with the auditory category regularities. In this way, both auditory category and visuomotor sequence regularities could benefit response to the visual target across training blocks and the task mimicked both SRT and SMART. In Block 7 the consistent auditory-category-to-location mapping was eliminated, but the visuomotor sequence continued to predict the visual target. Jointly, these first two conditions reveal whether incidental auditory category learning is influenced by the presence of an implicit visuomotor sequence that also predicts response to the visual target.

The remaining two conditions examined visuomotor sequence learning in the presence or absence of incidental auditory category learning. In the *Sequence, No Category* condition the 12-trial sequence aligned with the visual target locations across six training blocks. The sequence was eliminated by randomization in Block 7, and re-established in Block 8. This condition is similar to traditional SRT approaches, as described above. One difference was that auditory exemplars preceded the appearance of the visual target. In the *Sequence, No Category* condition, there was no category-to-location mapping and no within-category similarity; the five sounds preceding a trial were randomly selected across all four auditory categories. In this way, the presence of sounds preceding the visual target matched the conditions of the final *Sequence + Category (Sequence Violation)* condition, but there was no auditory category regularity aligned with behavior in the visuomotor task. In the final *Sequence + Category (Sequence Violation)* condition both the 12-trial visuomotor sequence and auditory category regularity predicted the visual target location in training blocks. In Block 7, the visuomotor sequence was randomized and the auditory-category-to-location information persisted. Together, these latter two conditions reveal whether visuomotor sequence learning is impacted by the presence of auditory categories simultaneously predictive of target location.

Results

Incidental auditory category learning occurs even in the presence of visuomotor sequences We first examined post-test performance in the *Sequence, No Category* condition for which only the visuomotor sequence regularity was available to support visuomotor task performance. Since there was no auditory category regularity aligned with visuomotor task performance in the primary task, we did not expect incidental auditory category learning (Gabay et al., 2015; Roark et al., 2021). Indeed, participants in the *Sequence, No Category* condition exhibited at-chance performance in overtly labeling the novel auditory category exemplars at post-test, t (25) = -.02, p = .97; Cohen's d = -.08.

We next examined the other three conditions to understand the impact of the presence of visuomotor sequence regularities on incidental auditory category learning. When auditory regularities were available, and aligned with the visuomotor task, there was robust evidence of generalization of incidental auditory category learning to overt labeling of novel category exemplars, with accuracy significantly above-chance (25%; Fig. 2a) for each condition. This was true when auditory regularities were present in isolation in the *Category, No Sequence* condition (M = .46, SE = .04, $t(24)$ = 4.82, p < .001; Cohen's d = .95) and also when visuomotor sequences were present in collaboration with the auditory category regularities in the *Category + Sequence (Category Violation)* condition (M = .36, SE = .03, $t(25)$ = 2.95, p = .006; Cohen's d =.55) and the *Sequence + Category (Sequence Violation)* condition (M = .40, SE = .04, $t(27)$ = 3.53, p < .001; Cohen's d =.68).

Notable with regard to the logic of the dual-task approach, a one-way ANOVA revealed that there was no significant difference in performance across these three conditions, F (2, 76) = 1.26, p = .28; η_p^2 = .03; BF = .29.[1] Incidental auditory category learning was not reduced in the presence of the additional visuomotor sequence regularity aligned with the same aspects of the primary visuomotor task.

The pattern of RTs in training is important in understanding these outcomes. In both the *Category, No Sequence* and *Category + Sequence (Category Violation)* conditions, the category-to-location mapping (but not the visuomotor sequence mapping) was eliminated in Block 7. In line with prior research (Gabay et al. 2015), we predicted that this would slow the RT to respond to the visual target if participants had incidentally learned auditory categories as predictors of the upcoming visual target location. Indeed, there was a significant slowing of RT when the category-to-location was destroyed in Block 7 (M = 435.4 ms, $S.E.$ = 11.9 ms) compared to Block 6 (M = 414.0 ms, $S.E.$ =15.1 ms), F (1, 49) = 10.11, p = .003; η_p^2 = .17. (Fig. 2c) with no difference in overall RT across conditions, F (1, 49) =.23, p=.63; η_p^2 = .004. There was a significant condition-by-block interaction, F (1, 48) = 5.11, p = .02; η_p^2 = .09, revealing a difference in the magnitude of the RT Cost across conditions. Whereas there was a significant RT Cost for auditory regularities in isolation (*Category, No Sequence*, t (24) = 3.522, p = .0017, Cohen's d =.74, M_{RTCost} = 36.57, SE_{RTCost} = 10.38), there was no RT Cost when visuomotor sequences were also present (*Category + Sequence (Category Violation)* condition, t (25) = .71, p = .47; Cohen's d =.14, M_{RTCost} = 6.17, SE_{RTCost} = 8.60).

The lack of RT Cost in the *Category + Sequence (Category Violation)* condition should not be taken as an indication that the incidental auditory category learning did not occur: participants in this condition were able to overtly label novel category exemplars with accuracy that significantly exceeded chance performance, and was on par with generalization of category knowledge among participants in the *Category, No Sequence* condition. Here, the lack of RT Cost is informed by the fact that participants had the advantage of a visuomotor sequence regularity that persisted in Block 7, even as the auditory category regularity was violated. Thus, the lack of a RT Cost, combined with generalization of incidental auditory category learning in the overt post-test in the *Category + Sequence (Category Violation)* condition is indicative of learning across *both* visuomotor sequence regularities and auditory category regularities.

Visuomotor sequence learning occurs even when simultaneous auditory category regularities can benefit task performance Visuomotor sequence learning was not abolished by the presence of another source of information that could benefit task performance. In conditions conveying a visuomotor sequence (*Sequence, No Category; Sequence + Category (Sequence Violation)*), RTs were slower when the visuomotor sequence was destroyed in Block 7 (M = 472.8 ms, SE = 15.5 ms) compared to Block 6 (M = 454.3 ms, SE = 18.7 ms), F (1,

[1] Because it is risky to accept the null hypothesis, we also calculated a Bayes factor (BF) for the main effect of group. The Bayes factor states the ratio between the evidence supporting the hypothesis relative to the null hypothesis (Dienes, 2011), such that a Bayes factor with a value of less than 1/3 indicates support for the null hypothesis. In contrast, a Bayes factor over 3 suggests that the analysis supports H1. Bayes factors were calculated using JASP – a free software for statistical analysis.

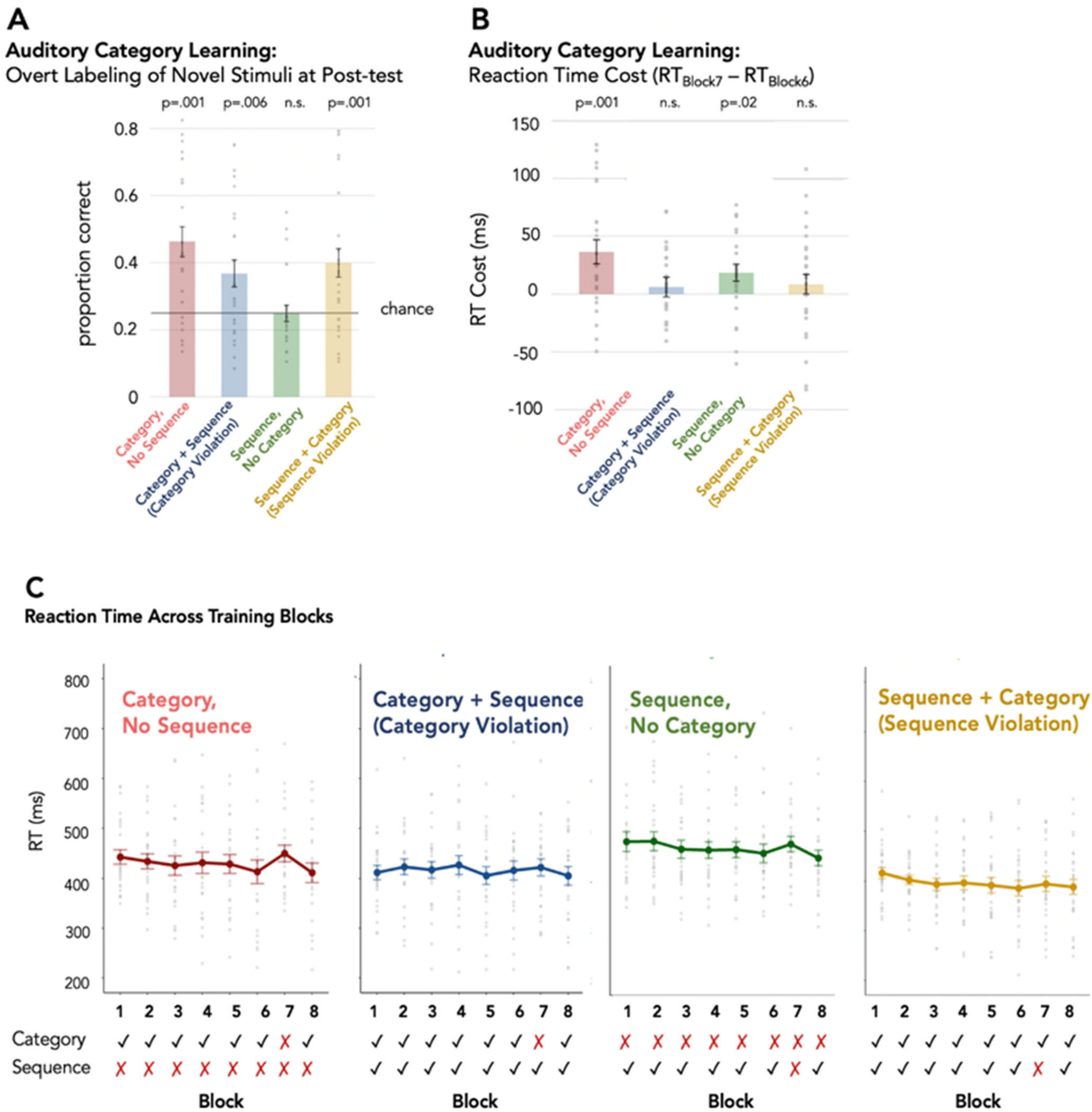

Fig. 2 Overview of learning across the four conditions. **a** Average accuracy in the post training overt categorization task across conditions. Note that no auditory category regularities were aligned with the primary task in the *Sequence, No Category* condition. All sounds categorized in the overt categorization task were novel category exemplars not experienced in training and therefore demanded generalization of learning. Chance-level performance is .25. **b** Average response time (RT) Cost (the difference in average RT across Blocks 6 and 7) across conditions provides a "covert" measure of category learning that does not require explicit labeling. **c** RT to detect the visual target as a function of block, presented across conditions. In all panels, error bars indicate standard error of the mean and grey dots indicate individual participant performance

51) = 5.7, p = .02; η_p^2 = .10 across conditions (Fig. 2c). There was no interaction of condition and block, F<1, indicating similar-magnitude RT Costs when only visuomotor sequences were available to support learning in the *Sequence, No Auditory Category* condition (M_{RTCost} = 18.56, SE_{RTCost} = 7.28) versus when auditory category regularities were also present in the *Sequence + Category (Sequence Violation)* condition (M_{RTCost} = 8.58, SE_{RTCost} = 8.54). Together, this indicates that visuomotor sequence learning can occur even while auditory category regularities are aligned to benefit task performance, but category learning did not sustain faster RTs when visuomotor sequences were violated (as visuomotor

sequences did upon category violations). Participants were faster in RT in the *Sequence + Category (Sequence Violation)* condition compared to the *Sequence, No Category* condition, $F\,(1, 51) = 11.07$, $p = .001$; $\eta_p^2 = .17$.

General discussion

Both incidental category learning and visuomotor sequence learning can be evoked in tasks that involve learning predictive relationships from input regularities without explicit instruction or knowledge of the regularities. But learning is not entirely passive, or feedback-free. Rather, the sensory regularities are aligned with actions and provide information with which to predict future action, the outcomes of which can provide a form of *internal feedback* from successful predictions. Capitalizing on the logic of dual-task conditions to examine whether simultaneous auditory category and visuomotor sequence regularities interact to facilitate or interfere with learning, we observe parallel learning of auditory category and visuomotor sequence regularities that jointly predict behavior in an ostensibly unrelated task. The degree of generalization of auditory category learning to novel exemplars, a benchmark of robust learning, was not impacted by simultaneous learning across visuomotor sequences signaling the same outcomes. Despite the surface commonalities across SRT and SMART tasks, and partly similar modalities, the simultaneous availability of visuomotor sequence and auditory category regularities did not compete to negatively impact learning.

At the broadest level, the results emphasize the importance of the alignment of an input regularity with the primary visuomotor task in supporting learning. Participants were not informed about the existence of auditory categories or visuomotor sequences, did not overtly search for dimensions diagnostic to these regularities, and did not receive explicit feedback. Notably, participants in each condition, including *Sequence, No Category*, experienced the same auditory category exemplars. But, in contrast to the other three conditions, category regularities in the *Sequence, No Category* condition did not align with behaviorally relevant demands of the visuomotor task. Overt category labeling performance in conditions in which auditory category regularities were aligned with behaviorally relevant actions and events (*Category, No Sequence*; *Category + Sequence (Category Violation); Sequence + Category (Sequence Violation)*) revealed robust generalization to novel sound exemplars, not present in the *Sequence, No Category* condition. Thus, we can conclude that the auditory categories were learned *incidentally* by virtue of their alignment with a task defined by other, here largely visuomotor, task demands.

Yet, participants in the *Sequence, No Category* condition had passive exposure to the same auditory category exemplars. Thus, the lack of auditory category learning in the *Sequence, No Category* condition demonstrates that the statistical learning of category input regularities observed in the other conditions was not entirely passive, or evoked via mere exposure. Instead, incidental category learning occurred when auditory category regularities aligned with behaviorally relevant outcomes (Roark et al., 2021) providing the "representational glue" with which to bind together acoustically variable exemplars that possess some underlying statistical regularity (Lim, Fiez & Holt, 2019). In this way, incidental learning of input regularities in the context of an ostensibly unrelated task can provide an "active assist" to passive statistical learning. Here, this conclusion is further corroborated by prior research demonstrating that the multidimensional auditory categories used in the present study are not learned across passive exposure (Emberson et al., 2013; Wade & Holt, 2005).

Crucial to the aims of this study, generalization of incidental auditory category learning across the three conditions (*Category, No Sequence, Category + Sequence (Category Violation, Sequence + Category (Sequence Violation)*) for which category input regularities aligned with the visuomotor task was not modulated by the presence of simultaneous visuomotor regularities also aligned with the task. Learning these two input regularities proceeded without interaction.

The pattern of behavior observed in training is important in understanding these outcomes. Notably, evidence from the training phase revealed no reliable RT Cost in destroying the alignment of auditory categories with the visuomotor task in the *Category + Sequence (Category Violation)* condition. In contrast, there was a RT Cost to when auditory categories were aligned with the visuomotor task, with no visuomotor sequence regularities in the *Category, No Sequence* condition. This is understood in relation to successful post-test generalization of incidental auditory category learning in both conditions. Participants in the *Category + Sequence (Category Violation)* condition appear to have relied upon the visuomotor sequence regularity to avoid suffering a RT Cost with misalignment of auditory categories to visuomotor task. Combined with successful auditory category learning at post-test, this indicates simultaneous learning of visuomotor sequence and auditory category input regularities. Inasmuch as generalization of auditory category learning was as successful at post-test for the *Category + Sequence (Category Violation)* and the *Category, No Sequence* conditions, it does not appear that the joint presence of the regularities had a negative impact on auditory category learning. In a similar manner, when visuomotor sequences were aligned with the visuomotor task in the *Sequence, No Auditory Category* and *Sequence + Category (Sequence Violation)* conditions we observed similar-magnitude RT Costs upon randomization of the sequence. This suggests that visuomotor sequence learning can occur even in the presence of auditory regularities.

Notably, the RT Cost was absent in the *Category + Sequence (Category Violation)* condition and differed significantly from the RT Cost in the *Category, No Sequence* condition. However, it was of similar magnitude in the *Sequence, No Auditory Category* and *Sequence + Category (Sequence Violation)* conditions. This pattern of results suggests an asymmetrical relationship between the SRT and SMART tasks; when auditory violations occur in the SMART task, the visual sequence from the SRT is sufficient to support learning; but when sequence violations occur, the presence of auditory regularities does not support learning. This may suggest a dominance of visual-motor information compared to auditory-visual or auditory-motor information, or an asymmetry in difficulty across tasks.

Response latencies were significantly faster in the *Sequence + Category (Sequence Violation)* condition compared to the *Sequence, No Category* condition. It is possible that converging auditory and visuomotor sequence regularities facilitated visuomotor task performance. However, there was no response latency advantage across *Category, No Sequence* and *Category + Sequence (Category Violation)* conditions. Here, it is important to recall that the *Sequence, No Category* condition included sounds to parallel the other conditions. The difference was that auditory category input regularities did not align with the visuomotor task in the *Sequence, No Category* condition; this misalignment may have slowed RTs. A recent study demonstrates that the presence of irrelevant sounds in the context of visuomotor sequence learning slows visuomotor task performance compared to a control condition with no sounds or a condition with task-aligned sounds (Robinson & Parker, 2021).

Learning in the context of dual input regularities is especially notable. Learning contexts in which both regularities aligned with behavior on the primary, visuomotor task were learned just as well as learning of either regularity in isolation. In both sequence learning and incidental auditory category learning participants used input regularities to guide actions and had the opportunity to use *internal feedback* derived from successfully making predictions about upcoming events, to guide behavior. Despite these similarities across the SRT and SMART tasks, learning visuomotor sequence regularities and learning auditory category regularities do not appear to draw upon a common pool of cognitive resources in a competitive manner (Kahneman, 1973; Posner & Petersen, 1990; Tombu & Jolicœur, 2003). In this regard our results are in line with previous findings showing that dual task (word learning and math tasks, but not a sentence processing task) presented via the auditory modality did not interfere with visuomotor sequence learning (Nemeth et al., 2011).

Furthermore, inspired by overshadowing designs in the conditioning literature (Kamin, 1968; Tennant & Bitterman, 1975), the present design allows for examination of learning when auditory category regularities are redundant with visuomotor sequence regularities (*Category + Sequence (Category Violation), Sequence + Category (Sequence Violation)* and thus potentially carry less informational value. Overshadowing, the observation that when joint information sources (X, Y; here, e.g., visuomotor sequence and auditory category regularities) predict the same outcome, subsequent response to either X or Y alone is less robust than had X and Y been learned in isolation (Kamin, 1968; Matzel et al., 1985). Thus, the redundant regularities in the two conditions mentioned above might be expected to hinder learning, as in overshadowing effects, because a fixed amount of associative strength is available and distributed across multiple cues (as, e.g., in the Rescorla-Wagner model, 1972). Instead, we observe no evidence of a reduction in learning that tracks with predictions from overshadowing.

The present data demonstrate that motor sequence knowledge and incidental auditory category knowledge can be acquired independently and simultaneous, even when each is predictive of the same outcome. This suggests the two learning challenges do not compete for the same cognitive resources. This relative independence might arise if conditional statistics (visuomotor sequence learning) and distributional statistics (incidental auditory category learning) rely on distinct learning mechanisms, as postulated by Thiessen et al. (2013). Or, the lack of interaction in the dual-task conditions may point to learning differences across perceptual modalities (Conway, 2020; Conway & Christiansen, 2005; Conway & Pisoni, 2008; Frost et al., 2019; Frost et al., 2015; Goschke, 1998; Goschke & Bolte, 2012). If so, we would expect that dual regularities within the same modality predict the same outcome to produce either facilitative (Beesley & Shanks, 2012) or interference effects (Endo & Takeda, 2004; Nixon, 2020), as in prior research. It is also possible that the two learning challenges were simply so easily learned that simultaneous learning by the same mechanism occurred, without cost. Alternatively, one might argue that learning performance for a single regularity is so good that additional information does not further facilitate learning. These latter two possibilities are unlikely to explain the present results; post-test categorization accuracy was far from ceiling.

In summary, young adult participants readily learned simultaneous visuomotor sequence and auditory category regularities incidentally when the regularities aligned with common demands in an ostensibly unrelated task. Moreover, learning outcomes did not differ substantially when these regularities occurred simultaneously, compared to learning either regularity in isolation. Thus, incidental learning across simultaneous statistical regularities – visuomotor sequence learning in the SRT task and incidental auditory category learning in the SMART task – does not appear to be in competition for a common pool of limited cognitive resources, or in competition for associative strength. Future work will be needed to determine whether this is possible due to inherent representational differences across auditory and visuomotor learning, or if distinct learning mechanisms are implicated. In either case, there is a need for deeper investigation into how learners balance the demands of accumulating information as it plays out simultaneously in multimodal input characterized by both regularity and variability, and aligned with active tasks.

Open practices statement The data and materials for the experiment are available via the Open Science Framework at https://osf.io/5ax97/. The experiments were not pre-registered.

Funding This research was supported by the Binational Scientific Foundation (2015227) and the National Science Foundation-Binational Scientific Foundation (2016867, NSF BCS1655126) grants to authors LLH and YG and by the National Science Foundation grant (BCS1655126, BCS1950054) to LLH and the Israel Science Foundation grant (734/22) to YG.

Declarations

Conflict of Interest The authors declare no competing interests.

References

Abrahamse, E. L., Jiménez, L., Verwey, W. B., & Clegg, B. A. (2010). Representing serial action and perception. *Psychonomic Bulletin & Review, 17*(5), 603–623.

Beesley, T., & Shanks, D. R. (2012). Investigating cue competition in contextual cuing of visual search. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(3), 709.

Chamizo, V., Aznar-Casanova, J., & Artigas, A. (2003). Human overshadowing in a virtual pool: Simple guidance is a good competitor against locale learning. *Learning and Motivation, 34*(3), 262–281.

Chapman, G. B., & Robbins, S. J. (1990). Cue interaction in human contingency judgment. *Memory & Cognition, 18*(5), 537–545.

Cleeremans, A. (1997). Sequence learning in a dual-stimulus setting. *Psychological Research, 60*(1), 72–86.

Conway, C. M. (2020). How does the brain learn environmental structure? Ten core principles for understanding the neurocognitive mechanisms of statistical learning. *Neuroscience & Biobehavioral Reviews, 112*, 279–299.

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(1), 24.

Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities: Pitting abstract against stimulus-specific representations. *Psychological Science, 17*(10), 905–912.

Conway, C. M., & Pisoni, D. B. (2008). Neurocognitive basis of implicit learning of sequential structure and its relation to language processing. *Annals of the New York Academy of Sciences, 1145*(1), 113–131.

Conway, C. M., Eghbalzad, L., Deocampo, J. A., Smith, G. N., Na, S., & King, T. Z. (2020). Distinct neural networks for detecting violations of adjacent versus nonadjacent sequential dependencies: An fMRI study. *Neurobiology of Learning and Memory, 169*, 107175.

Deocampo, J. A., King, T. Z., & Conway, C. M. (2019). Concurrent learning of adjacent and nonadjacent dependencies in visuo-spatial and visuo-verbal sequences. *Frontiers in Psychology, 10*, 1107.

Dezfouli, A., & Balleine, B. W. (2012). Habits, action sequences and reinforcement learning. *European Journal of Neuroscience, 35*(7), 1036–1051.

Dienes, Z. (2011). Bayesian versus orthodox statistics: Which side are you on?. *Perspectives on Psychological Science, 6*(3), 274–290.

Ellis, N. C. (2006). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics, 27*(2), 164–194.

Emberson, L. L., Liu, R., & Zevin, J. D. (2013). Is statistical learning constrained by lower level perceptual organization? *Cognition, 128*(1), 82–102.

Endo, N., & Takeda, Y. (2004). Selective learning of spatial configuration and object identity in visual search. *Perception & Psychophysics, 66*(2), 293–302.

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: the paradox of statistical learning. *Trends in Cognitive Sciences, 19*(3), 117–125.

Frost, R., Armstrong, B. C., & Christiansen, M. H. (2019). Statistical learning research: A critical review and possible new directions. *Psychological Bulletin, 145*(12), 1128.

Gabay, Y., & Holt, L. L. (2015). Incidental learning of sound categories is impaired in developmental dyslexia. *Cortex, 73*, 131–143.

Gabay, Y., Schiff, R., & Vakil, E. (2012a). Attentional requirements during acquisition and consolidation of a skill in normal readers and developmental dyslexics. *Neuropsychology, 26*(6), 744.

Gabay, Y., Schiff, R., & Vakil, E. (2012b). Dissociation between online and offline learning in developmental dyslexia. *Journal of Clinical and Experimental Neuropsychology, 34*(3), 279–288.

Gabay, Y., Dick, F. K., Zevin, J. D., & Holt, L. L. (2015). Incidental auditory category learning. *Journal of Experimental Psychology: Human Perception and Performance, 41*(4), 1124.

Gabay, Y., Karni, A., & Holt, L. L. (2019). *Overnight consolidation and retention of implicit and explicit knowledge of incidentally learned auditory categories*. Interdisciplinary Advances in Statistical Learning.

Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General, 117*(3), 227.

Goschke, T. (1998). Implicit learning of perceptual and motor sequences: Evidence for independent learning systems. In M. A. Stadler & P. A. Frensch (Eds.), *Handbook of implicit learning* (pp. 401–444). Sage Publications, Inc.

Goschke, T., & Bolte, A. (2012). On the modularity of implicit sequence learning: Independent acquisition of spatial, symbolic, and manual sequences. *Cognitive Psychology, 65*(2), 284–320.

Hazeltine, E., & Ivry, R. B. (2003). Neural structures that support implicit sequence learning. In Jiménez, L. (Ed.), *Attention and implicit learning* (pp. 71–107). John Benjamins Publishing Company. https://doi.org/10.1075/aicr.48.08haz

Hou, X. (2021). Learning two syntactic constructions simultaneously: a case of overshadowing. *Language and Cognition, 13*(3), 467–493.

Iao, L.-S., Roeser, J., Justice, L., & Jones, G. (2021). Concurrent visual learning of adjacent and nonadjacent dependencies in adults and children. *Developmental Psychology, 57*(5), 733.

Idemaru, K., & Holt, L. L. (2014). Specificity of dimension-based statistical learning in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 40*(3), 1009.

Janacsek, K., Fiser, J., & Nemeth, D. (2012). The best time to acquire new skills: Age-related differences in implicit sequence learning across the human lifespan. *Developmental Science, 15*(4), 496–505.

Janacsek, K., Shattuck, K. F., Tagarelli, K. M., Lum, J. A., Turkeltaub, P. E., & Ullman, M. T. (2020). Sequence learning in the human brain: A functional neuroanatomical meta-analysis of serial reaction time studies. *NeuroImage, 207*, 116387.

Jiménez, L., & Vázquez, G. A. (2011). Implicit sequence learning and contextual cueing do not compete for central cognitive resources. *Journal of Experimental Psychology: Human Perception and Performance, 37*(1), 222.

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Citeseer.

Kamin, L. J. (1968). *"Attention-like" processes in classical conditioning. Miami symposium on the prediction of behavior: Aversive stimulation* (pp. 9–31). University of Miami Press.

Lau, J. S.-H., Casale, M. B., & Pashler, H. (2020). Mitigating cue competition effects in human category learning. *Quarterly Journal of Experimental Psychology, 73*(7), 983–1003.

Leech, R., Holt, L. L., Devlin, J. T., & Dick, F. (2009). Expertise with artificial nonspeech sounds recruits speech-sensitive cortical regions. *Journal of Neuroscience, 29*(16), 5234–5239.

Lim, S. J., & Holt, L. L. (2011). Learning foreign sounds in an alien world: Videogame training improves non-native speech categorization. *Cognitive Science, 35*(7), 1390–1405.

Lim, S.-J., Lacerda, F., & Holt, L. L. (2015). Discovering functional units in continuous speech. *Journal of Experimental Psychology: Human Perception and Performance, 41*(4), 1139.

Lim, S.-J., Fiez, J. A., & Holt, L. L. (2019). Role of the striatum in incidental learning of sound categories. *Proceedings of the National Academy of Sciences, 116*(10), 4671–4680.

Liu, R., & Holt, L. L. (2011). Neural changes associated with nonspeech auditory category learning parallel those of speech category acquisition. *Journal of Cognitive Neuroscience, 23*(3), 683–698.

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review, 82*(4), 276.

Matsuzaka, Y., Picard, N., & Strick, P. L. (2007). Skill representation in the primary motor cortex after long-term practice. *Journal of Neurophysiology, 97*(2), 1819–1832.

Matzel, L. D., Schachtman, T. R., & Miller, R. R. (1985). Recovery of an overshadowed association achieved by extinction of the overshadowing stimulus. *Learning and Motivation, 16*(4), 398–412.

McLaren, I., & Dickinson, A. (1990). The conditioning connection. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 329*(1253), 179–186.

McLaren, I. P., Forrest, C., McLaren, R., Jones, F., Aitken, M., & Mackintosh, N. (2014). Associations and propositions: The case for a dual-process account of learning in humans. *Neurobiology of Learning and Memory, 108*, 185–195.

Mitchel, A. D., & Weiss, D. J. (2011). Learning across senses: cross-modal effects in multisensory statistical learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(5), 1081.

Nemeth, D., Janacsek, K., Csifcsak, G., Szvoboda, G., Howard Jr., J. H., & Howard, D. V. (2011). Interference between sentence processing and probabilistic implicit sequence learning. *PLoS One, 6*(3), e17577.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19*(1), 1–32.

Nixon, J. S. (2020). Of mice and men: Speech sound acquisition as discriminative learning from prediction error, not just statistical tracking. *Cognition, 197*, 104081.

Pavlov, I. P. (1927). *Conditioned Reflexes*. Oxford University Press.

Peigneux, P., Maquet, P., Meulemans, T., Destrebecqz, A., Laureys, S., Degueldre, C., ... Franck, G. (2000). Striatum forever, despite sequence learning variability: a random effect analysis of PET data. *Human Brain Mapping*, *10*(4), 179-194.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13*(1), 25–42.

Rauch, S. L., Whalen, P. J., Savage, C. R., Curran, T., Kendrick, A., Brown, H. D., ... Rosen, B. R. (1997). Striatal recruitment during an implicit sequence learning task as measured by functional magnetic resonance imaging. *Human Brain Mapping*, *5*(2), 124-132.

Reed, J., & Johnson, P. (1994). Assessing implicit learning with indirect tests: Determining what is learned about sequence structure. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(3), 585.

Rescorla, R. A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Current Research and Theory*, 64–99.

Roark, C. L., Lehet, M. I., Dick, F., & Holt, L. L. (2021). The representational glue for incidental category learning is alignment with task-relevant behavior. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 48(6), 769–784.

Robinson, C. W., & Parker, J. L. (2021). Tones slow down visuomotor responses in a visual-spatial task. *Acta Psychologica, 218*, 103336.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926–1928.

Schmidt, J. R., & De Houwer, J. (2019). Cue competition and incidental Learning:No blocking or overshadowing in the colour-word contingency learningprocedure without instructions to learn.*Collabra: Psychology, 5*(1), 15. https://doi.org/10.1525/collabra.236

Schwarb, H., & Schumacher, E. H. (2012). Generalized lessons about sequence learning from the study of the serial reaction time task. *Advances in Cognitive Psychology, 8*(2), 165.

Seitz, A. R., Protopapas, A., Tsushima, Y., Vlahou, E. L., Gori, S., Grossberg, S., & Watanabe, T. (2010). Unattended exposure to components of speech sounds yields same benefits as explicit auditory training. *Cognition, 115*(3), 435–443.

Szegedi-Hallgató, E., Janacsek, K., & Nemeth, D. (2019). Different levels of statistical learning-hidden potentials of sequence learning tasks. *PLoS One, 14*(9), e0221966.

Tennant, W., & Bitterman, M. (1975). Blocking and overshadowing in two species of fish. *Journal of Experimental Psychology: Animal Behavior Processes, 1*(1), 22.

Thiessen, E. D., Kronstein, A. T., & Hufnagle, D. G. (2013). The extraction and integration framework: A two-process account of statistical learning. *Psychological Bulletin, 139*(4), 792.

Tombu, M., & Jolicœur, P. (2003). A central capacity sharing model of dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance, 29*(1), 3.

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General, 134*(4), 552.

Vlahou, E. L., Protopapas, A., & Seitz, A. R. (2012). Implicit training of nonnative speech stimuli. *Journal of Experimental Psychology: General, 141*(2), 363.

Vuong, L. C., Meyer, A. S., & Christiansen, M. H. (2016). Concurrent statistical learning of adjacent and nonadjacent dependencies. *Language Learning, 66*(1), 8–30.

Wade, T., & Holt, L. L. (2005). Incidental categorization of spectrally complex non-invariant auditory stimuli in a computer game task. *The Journal of the Acoustical Society of America, 118*(4), 2618–2633.

Zhang, X., & Holt, L. L. (2018). Simultaneous tracking of coevolving distributional regularities in speech. *Journal of Experimental Psychology: Human Perception and Performance, 44*(11), 1760.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.