Published in final edited form as:

Cognition. 2016 December ; 157: 156–173. doi:10.1016/j.cognition.2016.09.002.

Incremental Implicit Learning of Bundles of Statistical Patterns

Ting Qian¹, T. Florian Jaeger^{2,3,4}, and Richard N. Aslin²

¹Department of Biomedical and Health Informatics, The Children's Hospital of Philadelphia

²Department of Brain and Cognitive Sciences, University of Rochester

³Department of Computer Science, University of Rochester

⁴Department of Linguistics, University of Rochester

Abstract

Forming an accurate representation of a task environment often takes place incrementally as the information relevant to learning the representation only unfolds over time. This incremental nature of learning poses an important problem: it is usually unclear whether a sequence of stimuli consists of only a single pattern, or multiple patterns that are spliced together. In the former case, the learner can directly use each observed stimulus to continuously revise its representation of the task environment. In the latter case, however, the learner must first parse the sequence of stimuli into different bundles, so as to not conflate the multiple patterns. We created a video-game statistical learning paradigm and investigated 1) whether learners without prior knowledge of the existence of multiple “stimulus bundles” — subsequences of stimuli that define locally coherent statistical patterns — could detect their presence in the input, and 2) whether learners are capable of constructing a rich representation that encodes the various statistical patterns associated with bundles. By comparing human learning behavior to the predictions of three computational models, we find evidence that learners can handle both tasks successfully. In addition, we discuss the underlying reasons for why the learning of stimulus bundles occurs even when such behavior may seem irrational.

Keywords

Bayesian nonparametric models; stimulus bundles; incremental learning; implicit learning; sequential input; memory

Introduction

One of the fundamental challenges in navigating the world is to guide our own behavior appropriately by forming a representation that captures the essential features of the task environment. Understanding how people construct this representation is a central problem in the study of learning and cognition. Importantly, in most real-world circumstances, this

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

learning process must rely on input that unfolds gradually over time. Several formalisms have been proposed to explain how such learning occurs. For example, under the framework of Bayesian belief updating, learners are assumed to represent the task environment as a probabilistic model and update their estimates of model parameters after each observation (e.g., Anderson, 1991; Sanborn, Griffiths, & Navarro, 2010). In connectionist and other associative theories, learners are assumed to represent the associative weights between the variables of the task environment (both observable and hidden), and revise them according to the degree to which their previous settings had correctly predicted a new observation (e.g., Love, Medin, & Gureckis, 2004; McClelland & Rumelhart, 1981; Sakamoto, Jones, & Love, 2008). Regardless of which broad category an incremental learning model falls into, a common assumption is that the task environment can be summarized by a *single* set of parameters. Under this view, learning is essentially a process of continuously *revising* this single set of parameters, as the average properties of the input will eventually converge onto the true properties of the task environment with more and more observations.

While this assumption holds true for many laboratory tasks that have been employed in learning experiments, many real-world situations may challenge its validity (see also Gallistel, Krishan, Liu, Miller, & Latham, 2014; Gershman, Blei, & Niv, 2010; Kleinschmidt & Jaeger, 2015; Kording, Tenenbaum, & Shadmehr, 2007; Yu & Cohen, 2008 for similar concerns). For example, consider a task where a naive learner observes a sequence of daily weather phenomena over the course of a year. For a knowledgeable learner, this sequence of stimuli will contain four subsequences that exhibit unique and localized patterns — a *spring* subsequence that mainly consists of sunny and warm days, a *summer* subsequence of hot days, a *fall* subsequence of rainy days and a *winter* subsequence of snowy days.¹ In other words, this sequence of weather patterns consists of multiple *stimulus bundles*, each of which consists of a subsequence of stimuli that are linked to the same underlying properties of the task environment (Qian & Aslin, 2014). If a learner ignores the existence of these stimulus bundles and keeps only a running average of the “daily weather”, the learned representation will be of poor predictive value as it does not converge onto the properties of any latent structure that we refer to as a *season*. Instead, if a learner is sensitive to the presence of stimulus bundles (or relatedly, the latent causes underlying them), they can form the representation that encodes four different average daily weather estimates, and by doing so, predict the weather given the current season with much better accuracy.

Although much research has been conducted on the topic of incremental learning, we know surprisingly little about whether people can successfully infer the presence of stimulus bundles in an incremental learning task, and more importantly, whether they can build a representation that captures the complexities of such a task environment. Here, we focus on the case where the stimuli in the same bundle collectively defines a *statistical pattern*. There are at least three challenges in learning such stimulus bundles. First, the latent state of a task environment, such as the notion of seasons in the above example, is not observable. As a

¹This example is purposefully simplified in at least two aspects. First, there are also systematic changes within a season that can be learned and predicted. Second, the learning of seasons is not a purely statistical learning problem, but one that depends on the knowledge of astrophysics. However, to a *naive* learner, the learning of seasons can be (almost) treated as a categorization problem where the concept of seasons emerges primarily from the observation of the statistical patterns in daily weather.

result, a learner must infer when one bundle ends and another begins from the sequentially observed input, possibly by evaluating any deviation in the observed input with respect to the consistency of the statistical pattern (Gebhart, Aslin, & Newport, 2009). With correctly inferred bundle boundaries, each observation can then be used to revise the belief about the latent properties of the environment that apply only in the corresponding state of that bundle (cf., Summerfield, Behrens, & Koechlin, 2011; Yu & Dayan, 2005). (For instance, the weather on a summer day will only be relevant to the learning of summer weather, but not the weather of other seasons.) We refer to this first challenge as **the problem of bundle boundaries**. This problem is also often referred to as *non-stationarity*, i.e., the presence of multiple statistical patterns in the input (see Aslin, 2014; Brown & Steyvers, 2009; Cohen, McClure, & Yu, 2007; Gallistel et al., 2014; Jaeger & Snider, 2013). Second, a task environment may enter the same state multiple times by returning to an old state, producing sequences of stimulus bundles that vary in content but share the same underlying characteristics (e.g., over multiple summer seasons, the exact sequences of weather patterns will differ but the general tendency of the summer weather will stay the same). Recognizing the latent state of the environment underlying a stimulus bundle and retrieving its learned representation from memory can greatly reduce the cost of relearning and predict the observations more accurately. However, doing so means that learners must be able to differentiate and identify the statistical patterns in sequential input, most likely based on only partial observations of a complete bundle. We refer to this second challenge as **the problem of bundle identity**. Finally, a task environment may contain *contextual* cues that are correlated with the *identities* of states underlying stimulus bundles (e.g., Freidin & Kacelnik, 2011; Gureckis & Love, 2009). For example, in our weather scenario, the presence of shorter periods of daylight serve as a contextual cue that the season is “winter” (note that a contextual cue is different from cues that are *causally* related to observed stimuli, see Speekenbrink & Shanks, 2010; Wasserman & Castro, 2005 for example). But, it is not necessarily clear to a naive learner what these cues are or how reliable they are in a task environment (cf., Griffiths & Tenenbaum, 2009). In addition, without learning the statistical patterns of stimulus bundles, contextual cues themselves are of little value since they are merely labels (e.g., knowing the next season is named “summer” does not help predict the weather unless one has already had experience with the summer season). We refer to this third challenge as **the problem of contextual cue validity**.

The goal of the present article is to investigate whether people are able to overcome these three challenges in an incremental learning task. Our research, as well as the perspectives outlined above, builds upon insights from several separate lines of research. Although not commonly viewed as such, the problems of stimulus bundle boundary, bundle identity, and contextual cue validity affect learning in all aspects of perception, motor control, and higher level cognition. Closely related to the first challenge — the problem of bundle boundaries — is the issue of change detection, which has been studied in the context of decision-making tasks (e.g., Behrens, Woolrich, Walton, & Rushworth, 2007; Boorman, Behrens, Woolrich, & Rushworth, 2009; Brown & Steyvers, 2009; Nassar, Wilson, Heasly, & Gold, 2010; Payzan-LeNestour & Bossaerts, 2011; Speekenbrink & Shanks, 2010; Wilson & Niv, 2012). In these studies, participants are either asked to make sequential choices among several alternatives with various reward rates, or required to make predictions about certain

variables of interest that in turn yield differential rewards. Either the reward rates associated with the alternatives or the variables themselves will change unpredictably, thus requiring participants to detect the changes and update their preferences. In our terminology, a series of trials with the same mapping between task variables and reward configuration would constitute a stimulus bundle, and after a change in this mapping, a new bundle begins. People detect such changes rather successfully and swiftly. However, there are two important limitations in relating the findings from this line of work to the problem of bundle boundaries in general. First, the change detection literature has almost exclusively focused on task environments with very simple statistical structures, such as changes in a Bernoulli distribution that specifies the reward configuration (Behrens et al., 2007), or changes in the mean of a Gaussian distribution that controls the properties of the stimuli (Nassar et al., 2010). It is unclear whether this ability to detect changes extends beyond those simple scenarios that have been investigated (see Gebhart et al., 2009; Kraljic, Samuel, & Brennan, 2008 for examples where learners facing complex problems fail to detect changes as quickly). Second, previous studies in this tradition have focused on changes that are signaled by immediate and explicit changes in reward. However, in everyday tasks, external reward may not be available and is often indirect and delayed. It is thus unclear how well learners will perform in an implicit learning task when they can only rely on the statistical pattern of the input itself as the primary means of detecting bundle boundaries.

The second challenge — the problem of bundle identity — concerns the ability of learners to see commonalities between various subsequences of input and to construct a representation that compactly encodes only the unique statistical patterns of the task environment (cf., Collins & Koechlin, 2012). A related behavior is reported in the animal conditioning literature, that after a period of behavioral extinction, the previously conditioned response of an animal can spontaneously recover (e.g., Sissons & Miller, 2009), be renewed (e.g., Bouton & King, 1983) or be reinstated (e.g., Rescorla & Heth, 1975). For animals to exhibit such behavior, they must be able to detect the boundary between bundles of conditioning trials and bundles of extinction trials, and represent the distinct causal contingencies associated with each type of bundle as separate and unique states of the task environment (cf., Gershman et al., 2010; Qian, Jaeger, & Aslin, 2012). Similarly in human learners, it has been shown that language users can adapt to the statistics of phonetic categories (Eisner & McQueen, 2005; Kraljic & Samuel, 2005, 2007; Norris, McQueen, & Cutler, 2003), words (Creel, Aslin, & Tanenhaus, 2008; Yildirim, Degen, Tanenhaus, & Jaeger, 2016), prosodic patterns (Kurumada, Brown, Bibyk, Pontillo, & Tanenhaus, 2012; Kurumada, Brown, & Tanenhaus, 2012), and syntactic structures (Fine, Jaeger, Farmer, & Qian, 2013; Kamide, 2012) of specific talkers and to retrieve these statistics at later times. The ability to adapt to talker-specific linguistic patterns indicates that listeners have constructed a representation that encodes the statistics of different talkers separately, where the linguistic stimuli produced by a different talker are associated with a unique bundle identity (for an account along these lines, see Kleinschmidt & Jaeger, 2015, 2016). However, language use is a task for which adult listeners can draw on rich prior knowledge about what constitutes a good cue to bundle identity: talker identities. Therefore, this line of work leaves open whether the ability to learn bundle identity extends beyond language processing, where such highly informative cues may be unavailable. In particular, it is an open question to what extent

people are able to learn and maintain bundle identities beyond language processing when there are multiple underlying statistical patterns occurring (and re-occurring) across many stimulus bundles in a single task session.

The third challenge — the problem of contextual cue validity — concerns the ability of learners to identify variables that are potentially helpful for learning bundle boundaries and encoding bundle identities (cf., Godden & Baddeley, 1975). While certain types of cues can be highly salient in signaling the presence of stimulus bundles, such as talker identity in speech perception as mentioned above (see also Gebhart et al., 2009; Weiss, Gerfen, & Mitchel, 2009), arbitrary cues in a novel task can be difficult to pick up on. In a typical visual search task, where participants need to locate a target quickly among a set of distractors, repeated exposure to visual cue-target pairs can greatly improve performance in familiar contexts, a phenomenon known as *contextual cuing* (see Chun, 2000, for a review). The finding of contextual cueing suggests that although inferring a valid contextual cue among many potential candidates seems like an intractable problem, the effect of contextual cues is usually highly effective and implicit, with participants being generally unaware of the contingency between the cues and the target stimuli (Chun & Jiang, 1998). We asked whether this implicit facilitation from contextual cues will be observed in the detection and identification of stimulus bundles, and if so, how contextual cues facilitate the learning of stimulus bundles.

Building on the findings of these previous studies, our efforts here are to integrate research questions concerning the detection of bundle boundaries, memory for previously encountered bundles, and the ability to discover and use contextual cues under the framework of incremental learning. The statistical patterns we investigate as well as the structure of the changes to these patterns are complex compared to most previous work, and the learning we investigate is incidental, rather than explicitly elicited. Further, the task we employ does not directly tie learning to immediate explicit changes in reward. Finally, we measure learning through changes in highly automatic behavior — the change in response times of targeted motor movements. To this end, we developed a computer game *Whack-The-Mole* as the basis for our two experiments. As implied by the name, the objective of the game is to hit the single active mole as fast as possible, as it pops up from one of four holes (see Fig. 1 for a screenshot²). During the course of a typical experimental session, a sequence of the mole's locations will be observed by the participant. This sequence of locations consists of multiple stimulus bundles, each of which exhibits a statistical pattern. Each statistical pattern is characterized by a unique categorical probability distribution that determines the likelihood that the mole will appear at each location (details about how these probability distributions are generated and how these distributions generate stimuli are described in the method section). Throughout this article, we will refer to the observable characteristics of a stimulus bundle as a “statistical pattern”, and the underlying generative entity as a “probability distribution”. Importantly, different stimulus bundles may share the same statistical pattern (after one or more intervening bundles with a different pattern), but the *observed* sequences of stimuli within each bundle will almost certainly differ as they are

²We would like to thank YouTube user *doodlekat1* for providing excellent video tutorials on how to draw cartoon animals. Cartoon animals featured in this game were created by following those tutorials.

dynamically generated from the corresponding probability distribution. Each stimulus bundle will end after some variable number of trials and be followed by a new stimulus bundle, with the constraint that two neighboring stimulus bundles always implement two different statistical patterns. Figure 2 illustrates the design of this underlying process.

Subjects in our experiments were instructed to play the game, without being made aware of the generative process underlying the game (i.e., Fig. 2), or the contextual cues. The only instruction subjects were given was to whack the mole as fast as they could. Like most everyday learning in the real world, the game provided no direct trial-by-trial feedback about the underlying stimulus bundles: the score displayed on the screen (see Figure 1) was not correlated with the probability distributions used on a given trial (see methods for detail). Nevertheless, there was an implicit benefit of detecting stimulus bundles and learning their corresponding statistical patterns: accurate beliefs about the underlying probability distribution of the current bundle would allow participants to anticipate the mole's location on the next trial and whack the mole faster *on average*. Moreover, there was also a benefit of storing mental representations of previously encountered statistical patterns: they would likely speed up the recognition of stimulus bundles that implemented a previously encountered statistical pattern of the task environment and thereby more rapidly reach asymptotic performance.

Overview of Experiments

Experiment 1: recurring statistical patterns with correlated visual cues—In Experiment 1, learners were exposed to an environment containing multiple bundles, where the statistical patterns associated with bundles reoccurred on an irregular basis. The background animals (see Fig. 1) served as a visual contextual cue for the statistical patterns of stimulus bundles in this experiment: each unique proportional distribution of the eight animals in the background was paired with an arbitrary parameter setting of a probability distribution that defines a statistical pattern. For instance, a contextual cue in the form of four dinosaurs, one rabbit, two hippopotamuses, and one snail may correspond to a 55-25-5-15% probability distribution, defined over the four locations where the mole can appear (as visualized in Figure 3). Importantly, this visual cue signaled the *identity* of the underlying distribution only; the number of each kind of background animal was *not* numerically correlated with the parameters of the underlying distribution (e.g., for a given participant, any configuration of the background animals could be associated with a 55-25-5-15% probability distribution). Figure 3 also illustrates that the pairing between background animals and probability distribution is always consistent within subjects. Deviating from most of prior work that explored the effect of contextual cues on visual search (e.g., Chun & Jiang, 1998, 1999; Mack & Eckstein, 2011), the spatial arrangement of the animals was not informative about bundle boundaries or identity. Instead, their positions were randomly reshuffled on every trial to make the relevant cue (i.e., the proportion of animals) as subtle as possible. This serves the goal of the experiment, which was to investigate whether human learners were able to detect bundle boundaries and represent the associated probability distributions when a signaling context — the distribution of the background animals — was available.

Experiment 2: recurring statistical patterns without correlated visual cues—

Experiment 2 differed from the first experiment in that the visual cues (i.e., the distribution of background animals) were uncorrelated with various statistical patterns. A random distribution of the background animals was selected at the beginning of a game and then applied on all trials, across all stimulus bundles to be encountered by that participant. The positions of these animals were still randomly reshuffled on each trial. The goal of this experiment was to explore whether the distribution of the mole's positions across trials alone would be sufficient for learning stimulus bundles in sequential input.

Method

Participants

A total of 34 students at the University of Rochester participated in the experiments, 21 in Experiment 1 and 13 in Experiment 2. Data from one subject in Experiment 1 were removed due to experimenter error. Subjects were paid \$7.50 each upon the completion of the task. All subjects reported normal or corrected-to-normal vision and demonstrated the ability to use a touchscreen.

Apparatus

Subjects interacted with the game on a 24-inch LCD touchscreen set at a resolution of 1680-by-1050 pixels. The angle and the height of the touchscreen were adjusted to the subject's preference prior to an experimental session. The game window occupied a 28-by-20-cm region in the center of the screen. Each subject was allowed to use only one finger from the dominant hand at all times during the experiment. A response area around each of the four mole locations was predefined and a correct response to the appearance of the mole consisted of a touch to that response area as soon as the mole (or part thereof) was visible on the screen. The dependent measure, reaction time (RT) in milliseconds (ms), was defined as the duration between the moment the mole was just visible on the screen and the moment it was hit.

Procedure

Learning in our experiments was implicit, as the experimenter simply told subjects to play a video game and made no mention of the stimulus bundles in the game. Subjects first completed 20 practice trials to become familiarized with the interface of the game. On these practice trials, participants were exposed to a statistical pattern generated from a uniform distribution, where the mole would appear at any hole equally likely.

After the practice trials, the game *randomly* generated the first categorical distribution that controlled the positions of the mole across trials (e.g., Categorical($\theta = 0.42, 0.17, 0.29, 0.12$)), along with a random distribution of background animals (i.e., the visual cue) in Experiment 1. (In Experiment 2, the distribution of background animals was fixed throughout an experiment to void the utility of such cues.) This categorical distribution was then implemented in the initial *stimulus bundle* of the experiment. The length of a stimulus bundle was sampled from a Poisson distribution with a mean of 40 trials. This categorical distribution remained in effect until the first stimulus bundle had reached its length. Then,

for each subsequent bundle, the game either implemented a completely novel probability distribution or changed back to one of the past distributions, with the constraint that it would never go back to the immediately preceding distribution (which would be a null change). The probability of changing into a familiar distribution c was:

$$p(\text{familiar distribution } c) = \frac{n_c}{N + \alpha} \quad (1)$$

where n_c is the number of stimulus bundles that had implemented the distribution c so far in an experiment (i.e., distribution frequency), and N the total number of stimulus bundles so far. The α parameter determines how frequently the environment changed into a novel probability distribution, when a new categorical distribution would be generated *randomly* (with the constraint that it was always different from all previous distributions) and become effective in the subsequent bundle. In Experiment 1, the new distribution was also accompanied by a new configuration of background animals. The probability of changing into a novel distribution was:

$$p(\text{novel structure}) = \frac{\alpha}{N + \alpha} \quad (2)$$

The α parameter in Equations (1) & (2) influences the generation of novel probability distributions. If its value is large, a change into a novel distribution is more likely, as compared to reusing previous distributions, and the environment as a whole will consist of more unique distributions. In both experiments, we set the value of α to 2, representing a modest tendency to create new distributions throughout the course of the experiment (the average number of unique distributions encountered by a learner was 7.6 in Experiment 1 and 8.0 in Experiment 2). Combined, Equations (1) & (2) produce a rich-get-richer effect — statistical patterns that have already reoccurred frequently in previous stimulus bundles are more likely to reoccur (cf., the *Chinese Restaurant Process*; Aldous, 1985).

An experimental session was divided into 5 blocks to reduce fatigue. In order not to abruptly truncate a stimulus bundle, the length of each block was also dynamically determined: When a subject reached the last trial of a stimulus bundle, if the total number of trials in the current block was already equal to or greater than 200, the block would be terminated. Otherwise, the block would continue with another stimulus bundle. Such a variable-length block design aimed to minimize statistical anomalies that would be otherwise introduced into the task environment. Information about previously encountered distributions was always preserved across blocks, so that each subsequent block had a wider range of familiar distributions to return to than its previous block. Therefore, subjects experienced mostly novel statistical patterns in the first block, and increasingly more familiar patterns (due to the reuse of existing probability distributions) in later blocks. Each subject completed a total of 5 blocks and had a 2-minute break between every two blocks.

A key point in the design of the current experiments is that the sequence of stimulus bundles and statistical patterns encountered by each subject were always dynamically generated

during a game. Therefore, although the overall experience was similar for all participants, each of them encountered a unique setting of stimulus bundles and statistical patterns (and contextual cues for participants in Experiment 1). Compared to a design where all participants would receive the same pre-determined bundle sequence, our design might appear to allow too much variation between subjects. However, learning was expected to follow the same general principles no matter how the underlying stimulus bundles of the environment were arranged. With the sequences of stimulus bundles varying between subjects, any effects of limited generality —those that would only apply to a particular bundle sequence— would likely be averaged out, and thus provide greater confidence in the results obtained.

Finally, participants received a global form of “feedback” consisting of a score strategically displayed in the top-right corner of the game window. This created a competitive environment where participants were encouraged to act as fast as they can by knowing a score was being kept, yet the specifics of the score were difficult to track on a trial-by-trial basis without severely degrading performance. More importantly, *on a trial-by-trial basis*, the score was only linearly related to RT, and thus provided no direct information about 1) the parametric settings of the probability distribution in the current stimulus bundle, 2) whether there was a change of stimulus bundle in recent trials, or 3) whether contextual cues were useful. Across trials, of course, the score would provide some form of feedback to participants — those who successfully learned the distribution of the mole’s location in the current stimulus bundle would obtain a higher score due to reduced RT, and those who quickly detected a change in stimulus bundle would also obtain a higher score due to a better adapted strategy. The use of feedback in this sense is then consistent with our goal of investigating *implicit learning*, because it would require participants to learn the complex probabilistic structures of the task environment in order to be highly efficient at this game.

Computational Modeling: Idealistic and Ideal Learners

To examine whether human learners successfully inferred the presence of multiple stimulus bundles in the input, we compared subject behavior to the predictions made by two “idealistic” learner models and one *ideal* learner model. The first two models are *idealistic* in the sense that they are intended to represent two opposite extremes of learning behavior in our experiment. Comparing human behavior to these models reveals whether the assumptions made by human learners are more similar to one of the extremes or the other. The *ideal* learner model, on the other hand, aims to capture the underlying assumptions of the task environment in this experiment as fully as possible. Comparing human behavior to the predictions of the ideal learner assesses the similarities between human behavior and the outcome of complex probabilistic inferences.

Note that although both the idealistic and ideal learners are implemented as Bayesian models, the assumptions that they make are unrelated to the Bayesian framework and can be implemented by other types of learning models. Here, we focus on the computational-level analysis of learner behavior in our experiment (Marr, 1982), which allows us to highlight *why* learners would be sensitive to the presence and identity of stimulus bundles in the input.

In the general discussion, we briefly discuss the consequences of our findings for the development of algorithmic-level implementations of bundle learning.

“Idealistic” Learners

The first idealistic learner model is the *global learner* model, which predicts the mole’s location on each trial by integrating evidence from the very first trial of an entire experiment to the current trial. The global learner is *idealistic* in the sense that it completely ignores the possibility of multiple stimulus bundles, and views all trials of an experiment as resulting from a single probability distribution. A single set of parameters is revised and updated as each new observation is made. Formally, this idealistic global learner is described by a simple Bayesian belief updating algorithm (based on a Dirichlet-Multinomial distribution), where the predicted probability of any location i on the t th trial is:

$$g_t(i) = \frac{n_{1:t-1}^i + 1}{t - 1 + 4} \quad (3)$$

In Equation (3), $g_t(i)$ denotes the predicted probability made by the global learner for location i on trial t , and $n_{1:t-1}^i$ denotes the number of times that the mole has appeared at location i in the previous $t - 1$ trials. The +1 in the numerator and +4 in the denominator are essentially placing a uniform Dirichlet prior on this Bayesian belief updating process, so that on the very first trial, the global learner predicts 0.25 for all locations in the game.

The other idealistic learner is the *local learner* model, which not only recognizes the existence of multiple stimulus bundles in the input, but also is (unrealistically) aware of the boundaries of different bundles. That is, for each trial, the local learner predicts the mole’s location by integrating evidence from only the first trial of the current stimulus bundle to the immediately preceding trial in that bundle. Formally, the probability of any location i on the t th trial predicted by the local learner is:

$$l_t(i) = \frac{n_{f_i:t-1}^i + 1}{t - 1 - f_i + 1 + 4} \quad (4)$$

where f_i indicates the trial number of the first trial of the current stimulus bundle, and $n_{f_i:t-1}^i$ refers to the number of times that the mole has appeared at location i in the current bundle. As can be seen from the similarities between Equation (3) and Equation (4), the local learner is essentially a simple Bayesian belief updating model that adapts its predictions to the statistical pattern of the current stimulus bundle (i.e., it assumes a separate Dirichlet-Multinomial distribution for each bundle). However, it is important to note that the local learner does not *remember* probability distributions learned from the statistical patterns of past stimulus bundles: the local learner follows Equation (4) regardless of whether a stimulus bundle implements a familiar or novel probability distribution.

Figure 4 illustrates the difference between the predictions made by the global learner and the local learner given the trial history of one participant. The predictions of the local learner exhibit much more variability and flexibility, because it essentially “resets” the frequency counts of the mole’s past locations at the beginning of a new stimulus bundle. In contrast, the predictions of the global learner converge onto some arbitrary distribution that results from averaging the mole’s locations over the entire experiment, while ignoring stimulus bundles.

The Ideal Learner

Intuition—The ideal learner model adopts assumptions that fall between the extreme ends represented by the two idealistic models. These assumptions are motivated by the central insight that, for a learner without perfect knowledge of the underlying latent structures, any sequential input can be subjectively perceived as non-stationary and organized in the form of potentially recurring stimulus bundles (Kleinschmidt & Jaeger, 2015; Qian & Aslin, 2014). The key feature of the ideal learner is then that it adopts an *inference* approach to discovering such bundle boundaries. Like the global learner, the ideal learner starts the learning process without knowing where the boundaries of stimulus bundles are. Unlike the global learner, it then uses the observed sequence of the mole’s locations throughout an experiment, incrementally, as a *cue* to infer the bundle boundaries (cf., Qian and Aslin 2014). For example, the ideal learner will consider a location sequence of 14321214123 more likely to form a single stimulus bundle than 111222333444, because the latter sequence consists of intuitive sub-patterns (i.e., repetitions of a single location). In our experiment, the actual location sequences are much noisier than 111222333444. The bundle boundaries inferred by the ideal learner are thus probabilistically distributed around the true bundle boundaries. While this inference approach introduces uncertainties about where the bundle boundaries are, it is more realistic than that of the local learner (which knows bundle boundaries ahead of learning) in capturing the nature of the task facing human learners, who presumably also need to infer the boundary locations from the stimulus input.

Besides an inference approach to the problem of bundle boundaries, the ideal learner is also capable of remembering the statistical patterns encountered in previous stimulus bundles. In the case of encountering a statistical pattern for the first time, the ideal learner predicts the mole’s location on the next trial by integrating evidence from the *estimated* first trial of that stimulus bundle to the current trial. In this aspect, the ideal learner behaves similarly to the local learner, except that the exact locations of the first trial of each bundle is unknown. In the case of encountering a familiar statistical pattern, the ideal learner can retrieve from its learned distributions the parameters that best describe the current statistical pattern. This strategy requires a reasonably accurate inference of the bundle boundaries, and serves as a solution to the problem of bundle identities, which is missing in both the global learner and the local learner.

The ideal learner implemented in this work falls short of the demands of the Whack-The-Mole experiment in one aspect: the use of visual contextual cues. The ideal learner can only infer the boundaries and identities of bundles based on the probability distributions of the mole’s locations, while ignoring the presence of background animals. Therefore, the ideal

learner model is strictly ideal for Experiment 1, but slightly ‘sub-ideal’ for Experiment 2. To work around this limitation, we compare the differences in the fit of the ideal learner to human behavior in both Experiments 1 and 2, from which we then investigate when visual contextual cues affected the learning of stimulus bundles.

Implementation: the ‘Hibachi Grill Process’ model—We implemented the ideal learner using the “Hibachi Grill Process” (HGP) mixture model, which was first proposed in Qian and Aslin (2014). The HGP mixture model is a Bayesian nonparametric model that can automatically infer the number of stimulus bundles in sequential input (from only one bundle to infinitely many), their boundaries, and the associated identities (i.e., probability distributions). The details of the implementation is described in the Appendix. The fundamental intuition is that the stimulus sequences in the Whack-The-Mole experimental exhibit a locally independent and identically distributed (i.i.d.) pattern. That is, once the boundaries of a stimulus bundle are known, its stimuli are independent of one another. As a result, the learning problem is to infer those boundaries and figure out which sequences of stimuli can be treated as i.i.d. This inference problem mimics the seating scenario at a restaurant serving Hibachi Grill: guests of the same group (bundle) will be seated at the same table (identity) and will share the dishes with other groups of guests; given a steady stream of customers, the server’s inference problem is to figure out which customers belong to the same group of friends. Qian and Aslin (2014) presents a more detailed discussion of this analogy, how the HGP mixture model compares to the Chinese Restaurant Process mixture model, and its application to other problems of learning and cognition.

Similar to the idealistic learners, the ideal learner also makes trial-by-trial predictions about the mole’s location using a stochastic search algorithm described in the Appendix. The predicted probabilities can then be compared to observed reaction times to evaluate the fit of the model.

Results and Discussion

If the human learners in our experiments successfully extracted stimulus bundles from sequential input, their behavior should be better described by the local learner than the global learner. On the other hand, if the global learner provided a better fit to subject data, it would indicate that human learners failed to overcome even the most basic challenge — the problem of bundle boundaries. At the same time, it is unrealistic to expect human learners to immediately detect the presence of stimulus bundles at the very beginning of an experiment; the evidence for bundles would gradually accumulate as learners were exposed to more and more systematic changes in the underlying probability distribution. Therefore, it is also likely that (1) the degree to which the predictions of the local learner fit human behavior may be relatively low at the beginning of an experimental session, and then increase over time, and (2) the ideal learner may provide a better fit to human behavior than the local learner in situations where perfect (or near perfect) awareness of bundle boundaries is clearly unrealistic. Finally, since neither of the idealistic learners, local or global, is capable of remembering probability distributions encountered in previous stimulus bundles, we expect the ideal learner to provide a better fit to RTs than the local learner if human learners successfully remembered previously learned probability distributions.

In behavioral experiments of learning and cognition, reaction times can usually be approximately fitted with a log-normal distribution due to a long tail (Ratcliff, 1993). At the same time, the probabilities predicted by the learner models are approximately normally distributed. To make those two variables more comparable, we performed a logarithmic transformation on subjects' reaction times. For each subject, log RTs that were more than two standard deviations away from the subject's mean log RT were considered outliers and removed from the analyses (which removed approximately 5% of all data).

Sensitivity to the presence of stimulus bundles

Figure 5 provides an overview of how well the predictions of the three models fit actual reaction times using the data of four subjects as examples. There are two qualitative patterns in this figure indicating that the global learner is an unlikely model of how participants learned. First, log RT is *consistently* negatively correlated with the probabilities predicted by the local learner and the ideal learner (Figs 5B and 5C; this negative correlation in fact holds for all subjects under these two models), but not with the probabilities predicted by the global learner (Fig 5A). The negative correlation between log RT and predicted probabilities is intuitive —when a model captures the underlying psychological process reasonably well, RT is expected to be fast if the location of the mole is predicted to have a high probability. However, the lack of a negative correlation, or in some cases, a weakly positive correlation (such as for subject 's9' in Figure 5), as observed in the global learner, is difficult to explain. The presence of such a positive correlation suggests that the global learner's strategy of ignoring bundles completely is an incorrect assumption about human behavior in our experiments.

Second, the predictions of the local learner and the ideal learner disperse much more evenly than the global learner. This is because, as we have previously shown in Figure 4, the predictions of the global learner always converge onto spurious probability distributions by averaging the mole's locations across bundle boundaries. These spurious distributions span a much narrower range (as they will inevitably converge onto the uniform distribution given enough bundles of stimuli) than the actual distributions associated with each stimulus bundle. Therefore, probability "clumps" are formed in the predictions of the global learner, where human RTs actually displayed a wide range of variability (Figure 5A). The observation that the predictions of the local learner and the ideal learner seem to better capture the gradient in human RTs further suggests that learners are likely aware of the potential existence of stimulus bundles.

These observations are quantitatively assessed by using both R^2 and Kolmogorov-Smirnov (KS) as the metrics of model fit. These two metrics assess two different aspects of model fit: variability (R^2) and functional form (KS). R^2 reflects the amount of variance in human RT data explained by a given learner model, obtained by regressing log RT per subject over the predictions of each model individually. The KS statistic evaluates the similarity in functional form between the empirical distribution function of log RTs and model predictions, obtained by first performing z-score transformation on the variables (as the KS statistic is sensitive to scale differences) and then conducting the two-sample Kolmogorov-Smirnov test per subject. A better fit is indicated by a higher R^2 value or a *lower* KS value. As we will see in

the rest of the article, although these two metrics never lead to different conclusions, each can be insensitive to the underlying patterns that exist in the data. Therefore, these two metrics are always presented together to ensure a comprehensive understanding of the results.

As shown in Figure 6, regardless of whether correlated visual cues were present, R^2 values are consistently lower for the global learner than the local learner ($p < 0.05$ in Experiment 1 and $p < 0.001$ in Experiment 2),³ as well as the ideal learner in Experiment 2 ($p < 0.001$; in Experiment 1, the difference is marginally significant $p = 0.1$). Using the Kolmogorov-Smirnov statistic as the measure of model fit, the global learner fares consistently poorly, producing a higher score than both the local learner and the ideal learner in both experiments ($p < 0.001$ for all comparisons).

Crucially, when both model fit metrics were computed *cumulatively* over the course of an experimental session (that is, over the 1st stimulus bundle, then over the 1st and 2nd stimulus bundles, and so on), as shown in Figure 7, it is striking to observe that the global learner and the local learner described human behavior equally well during the early stages of an experimental session (we will return to discuss the ideal learner in the next section): In Experiment 1, the fits of these two models are not distinguishable for the first 10 stimulus bundles according to R^2 (first 3 according to KS); In Experiment 2: the differences in fits are more pronounced but the same trend still holds. However, human learners behaved increasingly more like the local learner over the entire course of an experiment, presumably because they had started to become increasingly more sensitive to the stimulus bundles in the input after repeated exposure (i.e., an interaction effect between the number of stimulus bundles taken into consideration and the type of learner model, local vs global, is significant, $p < 0.01$ for both experiments and for both R^2 and KS). This result lends support to previous findings that suggest human learners become sensitive to multiple stimulus bundles in the input, if the sequential input consists of more than one switch of stimulus bundles (Zinszer & Weiss, 2013).

Taken together, these results confirm our qualitative observations and present an interestingly complex picture of how learners handle stimulus bundles — learners start with an assumption of treating all observations as coming from the same probability distribution, which is then gradually overridden as the evidence for multiple stimulus bundles accumulates.

Confusion around bundle boundaries

The comparisons of human learners' RTs to the predictions of the models strongly suggests that learners are sensitive to the presence of stimulus bundles. However, none of the analyses presented so far showed any difference in how the local learner and the ideal learner predict human RTs, judging by either the overall model fit in Figure 6 ($p > 0.1$ in all cases) or the incremental fit in Figure 7 (for both Experiments 1 and 2, $p > 0.1$ except when only

³These p-values are obtained via a linear mixed effects regression of the form $R^2 \sim (1|subject) + predictor$. This regression tests whether a given learner model fits human RT better than the other two, while accounting for certain subjects whose data may be inherently difficult to fit by any of the models. All p-values reported for the purpose of comparing R^2 or KS results were obtained in this manner.

considering the first stimulus bundle). This lack of meaningful distinctions is in fact *not* surprising — the analyses presented above aggregated data from all trials within a stimulus bundle, for which both the local learner and the ideal learner make similar predictions. The crucial question of whether learners *inferred* bundle boundaries, which highlights the difference between the local learner and the ideal learner, is largely ignored in the above analyses.

To focus on the question of whether learners inferred bundle boundaries, we examine the behavior of human learners during the first few trials after a switch in stimulus bundle (we shall refer to those trial as the *bundle-starting trials* hereafter). For the local learner, these bundle-starting trials will be immediately recognized and then used as evidence for estimating the probability distribution underlying the newly effective bundle. For the ideal learner, however, they are treated as a continuation of the current stimulus bundle until a sufficient number of such trials exhibit a consistent and distinctive statistical pattern, which triggers the recognition of a new stimulus bundle.

Therefore, when predicting the mole's locations on these bundle-starting trials, the local learner starts by assuming all locations are equally likely (reflecting the uniform prior in Equation (4)) and then gradually adjusts the distribution according to the unfolding sequence of observed locations. This strategy leads to an impoverishment of prediction variability on these bundle-starting trials: all first trials of a bundle will have a probability of 0.25; all second trials will have a probability of 0.4 if its location is the same as that of the first trial, 0.2 if different, and so on. As shown in Figure 8A, which uses the first 5 trials of all stimulus bundles in Experiment 1 as an example, the predictions of the local learner are highly “clumped” while RTs are clearly not. The ideal learner, in contrast, starts by using the probability distribution of the current bundle (“current” as far as the ideal learner is concerned) to predict the first few bundle-starting trials. Then, when the ideal learner realizes that those trials should have been considered as bundle-starting ones, it recognizes the new bundle, retroactively incorporates the evidence from those trials into the new bundle (but the “incorrect” predictions that have been made are not affected), and all future predictions will be based on the statistical pattern of the new bundle. As seen in Figure 8B, this strategy results in a much better match between the gradient of predicted probabilities and the gradient of reaction times for bundle-starting trials.

Moreover, when the model fit metrics are computed *cumulatively* over the trials of stimulus bundle, as illustrated by Figure 9, we observe the exact pattern one would expect from our hypothesis about bundle-starting trials: the ideal learner fits RTs the best at the beginning of a stimulus bundle, and is then gradually overpowered by the local learner as more trials are taken into consideration (i.e., an interaction effect between the number of trials taken into consideration and the type of learner model, ideal vs local, is significant, $p < 0.01$ for both experiments under R^2 and KS). This pattern is reflected in both R^2 and the KS statistic, but drastically more pronounced in the latter. We view the insensitivity of the R^2 measure as a classic example of the Anscombe's quartet problem (Anscombe, 1973) — statistical patterns that are clearly different can appear similar under summary statistics that are inadequate to encode their differences. For example, it is almost impossible to conclude that the local learner is a better model of human RTs based on Figure 8, which takes into consideration the

first 5 trials within a bundle, but its R^2 is in fact marginally higher than that of the ideal learner (see the left panel of Figure 9 where $X = 5$). As a result, we interpret the R^2 patterns as evidence supplementary to the strong effect indicated by the KS statistic.

The unsurprising finding here is that human learners were not as quick as the local learner in detecting and adapting to a new stimulus bundle, since they could not have possibly known the locations of bundle boundaries a priori as the local learner does. Crucially, on bundle-starting trials, RTs are predicted by the ideal learner model *better* than the local learner. This finding is critical because predictions on these bundle-starting trials are one of the primary distinctions between the ideal learner and the local learner. Observing a better fit for the ideal learner exactly due to this distinction then lends strong support to our hypothesis that human learners, like the ideal learner, have also adopted an inference approach in solving the problem of bundle boundaries.

Recognition of bundle identities

The model fit patterns over the course of an experimental session (Figure 7) and over the trials within a bundle (Figure 9) indicate that human learners are not only passively sensitive to the presence of stimulus bundles, but also actively inferring where one bundle ends and another begins. Since the characteristics of stimulus bundles are the statistical patterns according to which the within-bundle stimuli are sampled, the intuitive next step after having inferred the bundle boundaries is to learn the set of statistical patterns present in the input. As mentioned above, the statistical patterns in the Whack-The-Mole experiment can be summarized as straightforward categorical distributions. Crucially, there are fewer categorical distributions than stimulus bundles, meaning that the probability distributions implemented across bundles will reoccur over the course of an experiment. This recurrence of statistical patterns effectively assigns identities to stimulus bundles, and thus adds an additional dimension to the learning task where human learners could gradually become better at recognizing and adapting to a *familiar* statistical pattern.

Besides the inference approach toward bundle boundaries, the other primary distinction between the ideal learner and the local learner is exactly the capacity to remember probability distributions encountered in previous bundles. In spite of having the knowledge of exact bundle boundaries, the local learner treats each stimulus bundle as an independent entity, and thus re-estimates the probability distribution underlying a stimulus bundle from scratch every time. The ideal learner, since it is based upon an incremental clustering algorithm, can bootstrap from its existing belief on the parameters of a probability distribution (i.e., current estimates of the parameters of a cluster) once it has determined the statistical pattern is a familiar one (i.e., new input relevant to the cluster is observed). Figure 10 takes the 10 recurring appearances of a probability distribution experienced by a representative subject as an example (shown in columns in Panels A and B). The two models make different predictions about human RTs at the most probable (black) and the second most probable (orange) locations of that distribution. The design probabilities (i.e., the actual parameters of the probability distribution) for these two locations are 0.45 and 0.36, respectively. Since the probabilities predicted by the local learner are reset at the beginning of a bundle, a volatile pattern of estimates continues to be observed throughout all

encounters. For instance, at the 7th encounter, the local learner severely underestimates the probability of the second most probable location, presumably because it happens to be observed less frequently simply due to the stochastic nature of sampling. At the same time, although the predictions by the ideal learner still hovered around the true parameters with considerable noise, the overall pattern is much more stable: after the bundle-starting trials, the predicted probabilities quickly converged onto a configuration that is much more representative of the actual distribution, especially in the case of the 7th encounter. In fact, the predictions of the ideal learner have a negative correlation with the RTs on these trials (see Figure 10d, $\beta = -0.29$, $p < 0.001$), while the predictions of the local learner does not (see Figure 10c, $\beta = -0.01$, $p > 0.5$).

While the dynamic and probabilistic design of stimulus generation is a key to our experiments, it inadvertently resulted in only a small subset of stimulus bundles with underlying probability distributions similar to the above example in Figure 10. Essentially, the ideal stimulus bundle for this analysis should feature a probability distribution with a most probable location that is at the same time not too probable, so that there exists a close second most probable location. Such a relatively flat probability distribution is ideal for highlighting the effect on remembering previously encountered probability distributions: it takes the local learner a relatively long time to differentiate the top two locations *every time* it encounters the same probability distribution, while the ideal learner can do so more quickly in repeated encounters by retrieving the learned statistical pattern from memory. In contrast, extremely peaked distributions (where one location is overwhelmingly likely without a close second) offer the opportunity of observing the most probable location from the very beginning of a stimulus bundle at an extremely high rate. As a result, even though the local learner always learns from scratch in every encounter of the distribution, it can identify the most probable location so quickly that the resulting pattern mimics the asymptotic behavior of recognizing bundle identities by the ideal learner. It is even possible that human learners will be incentivized to behave more like the local learner when a stimulus bundle features a peaked distribution, as learning from scratch is already highly effective in reducing reaction times without incurring the potential overhead of this retrieval process.

As a result, we took two complementary approaches toward testing the hypothesis of recognizing bundle identities. The first method focuses on a small subset of stimulus bundles within each experiment that implemented probability distributions with the features that favor a memory-retrieval strategy toward the problem of bundle identity: the probability of the most probable location is at least 0.4, and the probability of the second most probable location is at least 0.3; in addition, we only extracted the top two most probable locations where the effect of memory will be potentially most prominent. The resulting subset represents about 10% of original data in each experiment. Like the example shown in Figure 10, these probability distributions, from the local learner's perspective, will remain ambiguous for a relatively long number of trials in the sense that there are two locations that can be the most probable location. However, to the ideal learner, such ambiguity can be quickly resolved once its statistical pattern and the associated probability distribution have been identified. Therefore, we expect the ideal learner to be the better model of human RTs than the local learner. This hypothesis is tested by means of linear mixed effects regression

(Baayen, Davidson, & Bates, 2008), where we regressed human RTs against the predictions of both the ideal learner and the local learner and included random intercepts per subject. In Experiment 1, removing the predictions of the ideal learner resulted in a larger reduction in model quality ($\chi^2(1) = 57.3, p < 0.001$) than removing those of the local learner ($\chi^2(1) = 40.2, p < 0.001$); in Experiment 2, removing the predictions of the ideal learner resulted a significant reduction in model quality ($\chi^2(1) = 15.3, p < 0.01$) while removing those of the local learner did not ($\chi^2(1) = 1.1, p > 0.1$). This supports the hypothesis that the ideal learner captures a critical property of incidental human learning over sequentially presented data.

The second approach filtered much less data, excluding only stimulus bundles where the probability of the most probable location is less than 0.4. The resulting subset represents about 80% of original data in each experiment. Given this noisier set of probability distributions, we hypothesize that the “peakiness” of a probability distribution, which is quantified by the entropy of a distribution, will play a crucial role: in stimulus bundles with highly peaked distributions, the local learner will capture the RT patterns very well, because these distributions represent cases where retrieving learned bundle identities would provide the least potential gain; in stimulus bundles with relatively flat distributions, the ideal learner will predict human RTs more accurately, because the noisy patterns of relatively flat distributions are better handled by the memory retrieval strategy.

This hypothesis is tested by adding an interaction between model predictions and the entropy of a probability distribution to the same linear mixed-effect regression analysis as used above. While both the ideal and local learners are correlated with human log RTs negatively on average (for the ideal learner, $\beta = -0.14, p < 0.001$ in Experiment 1 and $\beta = -0.17, p < 0.001$ in Experiment 2; for the local learner, $\beta = -0.26, p < 0.001$ in Experiment 1 and $\beta = -0.31, p < 0.001$ in Experiment 2), the interaction with the entropy of a distribution is negative for the ideal learner ($\beta = -0.096, p < 0.05$ in Experiment 1, $\beta = -0.14, p < 0.01$ in Experiment 2) but *positive* for the local learner ($\beta = 0.18, p < 0.001$ in Experiment 1, $\beta = 0.21, p < 0.001$ in Experiment 2). This positive interaction effect means that the local learner becomes less effective in predicting human RTs when a stimulus bundle is associated with a relatively flat probability distribution than a peaky one. On the contrary, the predictive power of the ideal learner remains strong, as the negative interaction means that the correlation between its predictions and human RTs only becomes increasingly more negative as the peakiness of a distribution decreases, suggesting that the ideal learner provides an even better fit to human learning behavior during exposure to such flat distributions.

In summary, both our analyses suggest that human learners are capable of retaining in memory information about the probability distributions learned in the past, learning bundle identity and thereby avoiding unnecessary relearning. Interestingly, the absence of contextual cues in Experiment 2 did not prevent the learning of bundle identities. This suggests that the statistical patterning of mole locations in a stimulus bundle alone provided enough information for learners to implicitly recognize and memorize its associated probability distribution. That is, even in tasks with latent statistical complexity that exceeded that of most previous experiments, human learners infer bundle identities and do so even in the absence of overt cues to bundle identity (see also Hikosaka et al., 1999 for similar learning in monkeys, though after much longer exposure).

Utility of contextual cues

One interesting aspect of the findings so far is that despite the absence of correlated cues, it appears that the behavior of learners in Experiment 2 was almost identical to those in Experiment 1: learners can rely on the statistical patterns of stimulus bundles alone to detect and remember the underlying probability distributions in both of our experiments. While the lack of visual contextual cues did not prevent learners from extracting stimulus bundles from the input, it is possible that within Experiment 1, visual cues of different properties may have influenced learning behavior. Note that in Experiment 1, the proportions of background animals (8 in total) serve as an arbitrary indexing cue to the identities of stimulus bundles. It is thus conceivable that certain configurations would be more salient than others. For example, a visual cue that consists of only 8 rabbits is presumably more salient than another one that consists of 2 rabbits, 2 dinosaurs, 2 hippopotamuses, and 2 snails.

This intuition can be captured by calculating the statistical *entropy* of the distribution of background animals. Visual cues with a more or less even distribution of animals will be assigned with a high entropy, reflecting a less salient cue. Visual cues with a highly skewed distribution of animals will be assigned with a low entropy, reflecting a more salient cue. We then split the data of Experiment 1 into a *high-entropy visual cue* group and a *low-entropy visual cue* group using the mean entropy of background animals as the mid-point.

Note that the ideal learner we have implemented is *not* equipped with the mechanism to make use of visual contextual cues. Therefore, we would expect human RTs in the high-entropy group to be better captured by the ideal learner than the local learner, where the cues are less salient so that human learners probably cannot use them effectively as well. In the low-entropy group, however, human learners should deviate from the predictions of the ideal learner slightly more, because the cues are more salient and easy to use. This hypothesis is examined in Figure 11, which shows how well each model predicts human RTs within a stimulus bundle, respectively in the cases of high and low entropy cues. Although the KS statistic indicates no qualitative difference between the high and low entropy bundles, the R^2 metric is in line with our hypothesis. The R^2 pattern under low-entropy visual cues (see the top-right panel of Figure 11) is similar to the pattern reported in Figure 9 — that the ideal learner captured a limited number of bundle-starting trials better than the local learner but not significantly so for the rest of the trials in a bundle. Then interestingly, the trend is almost entirely the opposite when the entropy of visual cues is high: the local learner accounted for the bundle-starting trials somewhat better than the ideal learner, but on average, human RTs are much better predicted by the ideal learner instead ($\beta = 0.013$, $p < 0.01$). Overall, the ideal learner is the better model for human RTs in stimulus bundles with high-entropy visual cues but the worse model in those with low-entropy visual cues (i.e., the interaction effect between the type of learner model and the entropy of visual cues is statistically significant $\beta = -0.027$, $p < 0.001$).

Figure 11 reveals several interesting ways in which contextual cues affect learning behavior in both Experiments 1 and 2. First, one might question why the unequivocal advantage of the ideal learner in the high-entropy group is absent in Experiment 2 (shown in the right panel of Figure 9), where the proportion of background animals remained the same across bundles of different probability distributions. After all, since only the statistical patterning of the mole's

location can be used to infer stimulus bundles, the ideal learner should have been the best model for human behavior. We postulate that this is because the background animals in Experiment 2 were effectively “confounds”: the lack of correlation with underlying probability distributions actually results in a noisier learning environment than that of the ideal learner. Human learners were burdened by the unhelpful background animals that never changed between bundles rather than learning to ignore them completely, resulting in decreased sensitivity to bundle boundaries and identities than the ideal learner. On the contrary, in the high-entropy group of Experiment 1, the background animals are still *perfectly correlated* cues for the underlying probability distributions despite their relatively low saliency. At a minimum, those cues would not interfere with the inference of stimulus bundles.

Second, it is intriguing to speculate how high saliency contextual cues led human learners to behave increasingly more like the local learner rather than the ideal learner. Why would a local memory-less learning strategy be more attractive to human learners when cues are seemingly more noticable and should have helped the retrieval of learned distributions? One possibility is that human learners try to find a reasonable balance between the costs and benefits of a probabilistic inference strategy: a probabilistic learning strategy is only worth the effort when the cue is highly uncertain and confusable, because adopting a probabilistic strategy then can potentially account for the error in retrieving learned probability distributions. When the contextual cue is highly salient, learners are relatively certain that a new stimulus bundle has begun. In that case, although learners may under-perform on the bundle-starting trials without retrieving learned distributions from memory, the cost can be quickly compensated for by a much simpler learning algorithm (i.e., the local learner) given the clear perception of bundle boundaries when the probability distribution changes. Such an interpretation is also consistent with the pattern that the ideal learner consistently provided a better fit to human RTs than the local learner *throughout* a stimulus bundle in the low saliency group, suggesting that low saliency cues call for the continuous use of a more complex learning strategy throughout a stimulus bundle. In this regard, the effects of contextual cues in our experiments represent a more complex scenario than the simple performance-boosting effect that has been previously noted in visual search tasks (cf., Brockmole & Henderson, 2006; Chun, 2000; Chun & Jiang, 1998, 1999; Jungé, Scholl, & Chun, 2007; Mack & Eckstein, 2011).

General Discussion

The present experiments suggest that human learners can successfully detect multiple stimulus bundles, recognize the underlying statistical pattern of a stimulus bundle, and retrieve from memory the knowledge about a familiar statistical pattern in an incremental learning task. Moreover, learners can accomplish these tasks without the use of contextual cues, suggesting that the statistical patterns of stimulus bundles, by themselves, are sufficient for inferring the boundaries and identities of bundles. When visual contextual cues are present, their effects depend on saliency: with highly salient contextual cues, which implies easy detection of bundle boundaries, learners’ behavior is more consistent with a memory-less heuristic strategy that focuses on quickly adapting to the local statistical pattern of the new bundle after a bundle switch occurs; with low-saliency cues, which implies difficulties

in detecting bundle boundaries, learners' behavior is better characterized by a probabilistic strategy that *infers* bundle boundaries and identities under higher levels of uncertainty.

Given the current findings, it remains intriguing to speculate about why human learners would learn stimulus bundles at all. Learning stimulus bundles is essentially a segmentation task where one tries to find the optimal boundaries in the input so that resulting subsequences of stimuli would exhibit locally coherent statistical patterns. In our experiment, we as experimenters knew that the observable stimuli had been generated by a multi-bundle probabilistic process (i.e., one that is depicted in Figure 2). Given that knowledge, the observed human behavior seemed “correct” or even “optimal” because it suggests that learners were successful in reverse-engineering the task environment. However, from the perspective of a naive learner, the quick adaptation to a “bundled world” is perhaps less justified — could the locally coherent statistical pattern arise just due to chance? If so, why learn it? In their seminal paper, Tversky and Kahneman (1974) argued that due to the (mis)belief in the law of small numbers, people will expect sequences of almost any length to exhibit the statistical characteristics of the larger population. Thus according to the theoretical perspectives of Tversky and Kahneman (1974), the learners in our experiment would in fact have acted “irrationally” by learning the bundles so quickly, because such learning *could* be essentially an unjustified interpretation of local statistical patterns. The global learner, which maintains a single aggregate statistical pattern, would be more rational, and possibly be better at predicting the mole's locations than the local learner in the long term.

The reason why human learners in our experiment did not behave like the global learner, in our opinion, is precisely because it is unrealistic for people (and animals alike) to harvest the benefits of learning on an infinitely long time horizon (although see Gureckis & Love, 2009 for why learners might act otherwise in some situations). That is, learners are more likely to adopt a learning strategy with a high utility under the current task demand. For the issue of detecting bundle boundaries in the Whack-The-Mole game, learners put a high value on temporally-local predictability, since it is directly relevant for the task at hand — being able to hit the mole as fast as possible on the next few trials is a goal of higher importance than on trials in the distant future. In more real-world tasks, temporally-local predictability matters more, as it could directly bring on significant biological consequences. For instance, deciding when to leave a depleting food source for a new one — a typical task of animal foraging (e.g., Dall, McNamara, & Cuthill, 1999; Hayden, Pearson, & Platt, 2011) — is another example where learning bundle boundaries (i.e., where one food source ends and another begins) is critically important and tremendously beneficial. According to this view, we believe that extracting stimulus bundles might be the default mode of learning and cognition, as it affords the agent a significant boost in the ability to better guide its actions for the current goals of a task (see Kleinschmidt & Jaeger, 2015; Qian et al., 2012, for similar arguments). Two of the authors argued in a previous study that learning stimulus bundles can even occur regardless of whether it actually recovers the “correct” generative model of the world: the perception of streak shooting in basketball, although not supported by actual statistics (Bar-Eli, Avugos, Raab, & Bareli, 2006; Gilovich, Vallone, & Tversky, 1985), can be understood as the result of discovering stimulus bundles in a sequence of basketball shots (Qian & Aslin, 2014). In general, we believe that understanding the learning

of stimulus bundles from the perspective of increased predictive utility can provide useful insights into human behavior in various task domains beyond implicit statistical learning.

In a similar vein, learners behaved more like the highly sophisticated ideal learner than the local learner when contextual cues are uncertain, because the benefits of retrieving learned distributions from memory are not overwhelmingly important when a new bundle has been clearly indicated by high-saliency cues. That is, other than handling the bundle-starting trials, learners can afford to use a simplistic strategy like the local learner to adapt to the current statistical pattern rather than incurring the cost of retrieving learned representations from memory. Relatedly, learners generally ignore multiple stimulus bundles in artificial grammar learning studies unless there is longer exposure to later bundles (Gebhart et al., 2009), or more than one appearance of familiar bundles (Zinszer & Weiss, 2013). Overall, these results suggest that the representation of recurring statistical patterns, or the use of such learned representations, could be driven by the utility as defined by task demand.

Algorithmic-level implementations of bundle learning

Despite its ability to derive trial-by-trial predictions, the ideal learner model is still a computational-level account for learners' behavior (i.e., an iterative batch algorithm). It is unlikely that human learners have actually performed such complex probabilistic inference over the parameters of the HGP mixture model during the experiment. The ideal learner also departs from the incremental nature of the current experiments in one key aspect: in order to predict the mole's location on the t th trial, the ideal learner performs repeated Monte Carlo sampling given all previous $t - 1$ trials in order to estimate bundle boundaries and the associated probability distributions. This is arguably an unrealistic description of how human learners discover stimulus bundles, as both the complexity of computation and the memory required for such inference can be overwhelming.

What are then the precise psychological mechanisms underlying the ability of human learners to learn multiple bundles in an incremental task? One direction is to explore computational methods that approximate the inferences assumed by the ideal learner, but are computationally simpler and cognitively more plausible (Gershman et al., 2010; Griffiths, Vul, & Sanborn, 2012; Lloyd & Leslie, 2013; Sanborn et al., 2010). Another direction is to investigate whether well-established psychological accounts of how recent experience affects subsequent motor actions could account for the type of bundle-learning behavior observed here. One such account is priming, or spatial priming in this particular case (e.g., Maljkovic & Nakayama, 1996). For example, a simple spatial priming model for the Whack-The-Mole task would hold that recent execution of a motor movement (hitting the mole at a specific hole location) would result in increased *activation* for that motor plan and thus facilitate subsequent execution of that movement. Such an account is consistent with the observation that when a bundle switch results in a drastically different probability distribution, reaction times at the most probable locations are initially slower (due to lower activation levels of corresponding movement plans under a priming account), and gradually become faster as more instances of those locations are observed. That is, a priming account may explain the effect that we refer to as **confusion around bundle boundaries**.

To shed more light onto this approach, we created simple priming models where the activation of each hole location decays exponentially as a function of the number of intervening trials. To simplify the testing procedure, we fixed the decay rate at 1 and instead varied the size of a “priming window”, which limits the number of previous trials that activation values are calculated from. (Equivalently, we could also use an unlimited priming window and vary the decay rate. A large decay rate will then have a similar effect to using a small priming window, and vice versa.) Four different priming window sizes were tested, ranging from only 1 trial to 4 trials, since Cleeremans and McClelland (1991) have shown that priming effects in serial reaction times are generally limited to a window size of 4. These models can also be construed as variations of the global learner, with exponentially weighted history observations and limited memory. We then examined how well these priming models can predict human RTs on the bundle-starting trials (the first 5 trials of a bundle) in our two experiments, as compared to the ideal learner. When the priming window was only 1 trial, the resulting model performed much worse than the ideal learner in predicting RTs (the priming account with a window size of 1 has a significantly higher KS statistic than the ideal learner: $t = 0.28$, $p < 0.0001$ in Experiment 1, $t = 0.25$, $p < 0.0001$ in Experiment 2). As the window size increases to 4, the priming model became increasingly more predictive of human RTs on bundle-starting trials, but its KS statistic remained significantly higher, albeit to a lesser and lesser extent, than the ideal learner ($p < 0.01$ in all cases). These preliminary results indicate that how well a priming account can explain the effect of confusion around bundle boundaries depends on the size of the priming window (or equivalently, the activation decay rate) selected by the modeler (i.e., a larger window size yields a better fit). On the one hand, we view this dependency on the window size as being potentially problematic, because we are not aware of any principled reason to favor one size over another *a priori*. The optimal size is likely task-dependent as a window size coinciding with the average bundle length will presumably be the best choice. On the other hand, this dependency points to the possibility that human learners may also have to learn and adjust the optimal “priming window” during the course of learning. Such a dynamic priming hypothesis is then in line with how our ideal learner infers the bundle boundary at the computational level. In other words, the *inference* of bundle boundaries seems to be a crucial aspect of human learning behavior in our studies.

In addition, while our current set of experiments were not specifically designed to tease apart the predictions of a priming account and our ideal learner, we suspect that a priming account will also encounter difficulties in predicting the bundle-identity effect without incorporating a mechanism that mimics the ideal learner’s ability to store distinct statistical patterns. Consider, for example, that only three of the four mole locations have been observed for a long stretch of trials, and then the mole pops out of the fourth location. In a priming account the only effect this should have is that RTs at this location should now be facilitated in subsequent trials (and relatively, a *minor* reduction in activation at all other locations). However, for the ideal learner this trial is highly informative: it signals that a new bundle may have started, which may implement a familiar probability distribution, and that a *rapid* readjustment in the probability estimates of all locations is necessary. In fact, an instance of this behavior is seen in the **recognition of bundle identities** effect reported in Figure 10. That is, human learners can often rapidly converge onto the correct probability

estimates of the locations in a familiar distribution, after a new bundle has been detected. The ideal learner explains this behavior by conditioning the estimation of probabilities on the *identity* of bundles. That is, the ideal learner can overcome the influence of the intervening trials between two bundles of the same identity (i.e., the same underlying distribution), because once the identity of a bundle is recognized, the intervening trials become irrelevant for the purpose of adaptation to the current bundle. However, an account in terms of priming, regardless of its window size or decay rate, would assign larger weights to those intervening trials because they are closer to the current trial and there is no notion of familiar bundles. In future research, we plan to specifically test how much influence the intervening trials between bundles of the same identity would have on RTs, since the ideal learner and a priming account would have sharply contrasting predictions in this aspect of bundle learning behavior.

Finally, a priming account needs to incorporate at least one non-trivial mechanism to account for another important aspect of our findings: the effect of contextual cues. We have observed that a more salient combination of background animals encouraged learners to gradually behave *more* like the local learner than the ideal learner. As mentioned above, this result is attributed to how contextual cues influence the complexity of the learning task — with a highly salient cue, the overall task complexity is lower, which makes it more likely for human learners to avoid probabilistic inferences (since the benefit is rather limited) but focus on local statistics. For a priming model to explain this effect, it would need to incorporate a mechanism that conditionally shortens its priming window (i.e., increases the influence of local trials) when contextual cues are more salient.

In summary, a priming-based account can become a plausible algorithmic-level theory of bundle-learning behavior, as long as it incorporates mechanisms to deal with the *a priori* unknown bundle lengths, the recognition of bundle identities, and the recurrence of familiar bundle identities. Further investigations of these topics will continue to enrich our understanding of the learning behavior of humans and animals in an inherently dynamic and seemingly unpredictable world.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by National Institutes of Health Grants HD-037082 (to R.N.A.), HD-067250 (to Daniel J. Weiss and R.N.A.), HD-075797 (to T.F.J.), and NSF CAREER IIS-1150028 (to T.F.J.).

References

- Aldous, DJ. Exchangeability and related topics. In: Hennequin, P., editor. *École d'Été de probabilités de saint-flour xiii* à T 1983. Berlin: Springer Berlin Heidelberg; 1985. p. 1-198.
- Anderson JR. The adaptive nature of human categorization. *Psychological Review*. 1991; 98(3):409–429.
- Anscombe FJ. Graphs in statistical analysis. *American Statistician*. 1973; 27(1):17–21.

- Aslin RN. Infant Learning: Historical, Conceptual, and Methodological Challenges. *Infancy*. 2014; 19(1):2–27. [PubMed: 26568702]
- Baayen RH, Davidson DJ, Bates DM. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*. 2008; 59(4):390–412.
- Bar-Eli M, Avugos S, Raab M, Bareli M. Twenty years of “hot hand” research: Review and critique. *Psychology of Sport and Exercise*. 2006; 7(6):525–553.
- Behrens TEJ, Woolrich MW, Walton ME, Rushworth MFS. Learning the value of information in an uncertain world. *Nature Neuroscience*. 2007; 10(9):1214–21. [PubMed: 17676057]
- Boorman ED, Behrens TEJ, Woolrich MW, Rushworth MFS. How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron*. 2009; 62(5): 733–43. [PubMed: 19524531]
- Bouton ME, King DA. Contextual control of the extinction of conditioned fear: tests for the associative value of the context. *Journal of Experimental Psychology: Animal Behavior Processes*. 1983; 9(3): 248–65. [PubMed: 6886630]
- Brockmole JR, Henderson JM. Recognition and attention guidance during contextual cueing in real-world scenes: evidence from eye movements. *Quarterly Journal of Experimental Psychology*. 2006; 59(7):1177–87.
- Brown SD, Steyvers M. Detecting and predicting changes. *Cognitive Psychology*. 2009; 58(1):49–67. [PubMed: 18976988]
- Chun MM. Contextual cueing of visual attention. *Trends in Cognitive Sciences*. 2000; 4(5):170–178. [PubMed: 10782102]
- Chun MM, Jiang Y. Contextual cueing: implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*. 1998; 36(1):28–71. [PubMed: 9679076]
- Chun MM, Jiang Y. Top-Down Attentional Guidance Based on Implicit Learning of Visual Covariation. *Psychological Science*. 1999; 10(4):360–365.
- Cleeremans A, McClelland JL. Learning the structure of event sequences. *Journal of Experimental Psychology: General*. 1991; 120(3):235–253. [PubMed: 1836490]
- Cohen JD, McClure SM, Yu AJ. Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*. 2007; 362(1481):933–42. [PubMed: 17395573]
- Collins A, Koechlin E. Reasoning, learning, and creativity: frontal lobe function and human decision-making. *PLoS Biology*. 2012; 10(3):e1001293. [PubMed: 22479152]
- Creel SC, Aslin RN, Tanenhaus MK. Heeding the voice of experience: The role of talker variation in lexical access. *Cognition*. 2008; 106(2):633–664. [PubMed: 17507006]
- Dall S, Mcnamara J, Cuthill I. Interruptions to foraging and learning in a changing environment. *Animal Behavior*. 1999; 57(1):233–241.
- Eisner F, McQueen JM. The specificity of perceptual learning in speech processing. *Perception & Psychophysics*. 2005; 67(2):224–238. [PubMed: 15971687]
- Fine AB, Jaeger TF, Farmer TA, Qian T. Rapid Expectation Adaptation during Syntactic Comprehension. *PLoS ONE*. 2013; 8(10):e77661. [PubMed: 24204909]
- Freidin E, Kacelnik A. Rational Choice, Context Dependence, and the Value of Information in European Starlings (*Sturnus vulgaris*). *Science*. 2011; 334(6058):1000–1002. [PubMed: 22096203]
- Gallistel CR, Krishan M, Liu Y, Miller R, Latham PE. The perception of probability. *Psychological Review*. 2014; 121(1):96–123. [PubMed: 24490790]
- Gebhart AL, Aslin RN, Newport EL. Changing Structures in Midstream: Learning Along the Statistical Garden Path. *Cognitive Science*. 2009; 33(6):1087–1116. [PubMed: 20574548]
- Gershman SJ, Blei DM, Niv Y. Context, learning, and extinction. *Psychological Review*. 2010; 117(1): 197–209. [PubMed: 20063968]
- Gilovich T, Vallone R, Tversky A. The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*. 1985; 17(3):295–314.
- Godden DR, Baddeley AD. Context-Dependent Memory in Two Natural Environments: on Land and Underwater. *British Journal of Psychology*. 1975; 66(3):325–331.

- Griffiths TL, Tenenbaum JB. Theory-based causal induction. *Psychological review*. 2009; 116(4):661–716. [PubMed: 19839681]
- Griffiths TL, Vul E, Sanborn aN. Bridging Levels of Analysis for Probabilistic Models of Cognition. *Current Directions in Psychological Science*. 2012; 21(4):263–268.
- Gureckis TM, Love BC. Short-term gains, long-term pains: how cues about state aid learning in dynamic environments. *Cognition*. 2009; 113(3):293–313. [PubMed: 19427635]
- Hayden BY, Pearson JM, Platt ML. Neuronal basis of sequential foraging decisions in a patchy environment. *Nature Neuroscience*. 2011; 14(7):933–9. [PubMed: 21642973]
- Hikosaka O, Nakahara H, Rand MK, Sakai K, Lu X, Nakamura K, Doya K. Parallel neural networks for learning sequential procedures. *Trends in Neurosciences*. 1999; 22(10):464–471. [PubMed: 10481194]
- Jaeger TF, Snider NE. Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*. 2013; 127(1):57–83. [PubMed: 23354056]
- Jungé, Ja; Scholl, BJ.; Chun, MM. How is spatial context learning integrated over signal versus noise? A primacy effect in contextual cueing. *Visual Cognition*. 2007; 15(1):1–11. [PubMed: 18725966]
- Kamide Y. Learning individual talkers' structural preferences. *Cognition*. 2012; 124(1):66–71. [PubMed: 22498776]
- Kleinschmidt DF, Jaeger TF. Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review*. 2015; 2(122):148–203.
- Kleinschmidt DF, Jaeger TF. Re-examining selective adaptation: Fatiguing feature detectors, or distributional learning? *Psychonomic Bulletin & Review*. 2016; 3(23):678–691. DOI: 10.3758/s13423-015-0943-z
- Kording KP, Tenenbaum JB, Shadmehr R. The dynamics of memory as a consequence of optimal adaptation to a changing body. *Nature neuroscience*. 2007; 10(6):779–86. DOI: 10.1038/nn1901 [PubMed: 17496891]
- Kraljic T, Samuel AG. Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology*. 2005; 51:141–178. [PubMed: 16095588]
- Kraljic T, Samuel AG. Perceptual adjustments to multiple speakers. *Journal of Memory and Language*. 2007; 56:1–15.
- Kraljic T, Samuel AG, Brennan SE. First impressions and last resorts: how listeners adjust to speaker variability. *Psychol Sci a J Am Psychol Soc/ APS*. 2008; 19(4):332–8.
- Kurumada, C.; Brown, M.; Bibyk, S.; Pontillo, D.; Tanenhaus, M. Rapid adaptaton in online pragmatic interpretation of contrastive prosody. *Proceedings of the 36th annual conference of the cognitive science society*; Austin, TX: Cognitive Science Society; 2012. p. 791-796.
- Kurumada, C.; Brown, M.; Tanenhaus, MK. Pragmatic interpretation of contrastive prosody: It looks like speech adaptation. *Proceedings of the 34th annual conference of the cognitive science society*; Austin, TX: Cognitive Science Society; 2012. p. 647-652.
- Lloyd K, Leslie D. Context-dependent decision-making: a simple Bayesian model. *Journal of The Royal Society Interface*. 2013; (February)
- Love BC, Medin DL, Gureckis TM. SUSTAIN: a network model of category learning. *Psychological review*. 2004; 111(2):309–32. [PubMed: 15065912]
- Mack SC, Eckstein MP. Object co-occurrence serves as a contextual cue to guide and facilitate visual search in a natural viewing environment. *Journal of Vision*. 2011; 11(9):1–16.
- Maljkovic V, Nakayama K. Priming of pop-out: II. the role of position. *Perception & Psychophysics*. 1996; 58(7):977–991. [PubMed: 8920835]
- Marr, D. *Vision*. W.H. Freeman; 1982.
- McClelland J, Rumelhart D. An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*. 1981; 88(5):375–407.
- Nassar MR, Wilson RC, Heasley B, Gold JI. An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *Journal of Neuroscience*. 2010; 30(37):12366–78. [PubMed: 20844132]

- Norris D, McQueen JM, Cutler A. Perceptual learning in speech. *Cognitive Psychology*. 2003; 47:204–238. [PubMed: 12948518]
- Payzan-LeNestour E, Bossaerts P. Risk, unexpected uncertainty, and estimation uncertainty: Bayesian learning in unstable settings. *PLoS computational biology*. 2011; 7(1):e1001048. [PubMed: 21283774]
- Qian T, Aslin RN. Learning bundles of stimuli renders stimulus order as a cue, not a confound. *Proceedings of the National Academy of Sciences*. 2014; 111(40):14400–14405.
- Qian T, Jaeger TF, Aslin RN. Learning to Represent a Multi-Context Environment: More than Detecting Changes. *Frontiers in Psychology*. 2012; 3(July):228. [PubMed: 22833727]
- Ratcliff R. Methods for dealing with reaction time outliers. *Psychological Bulletin*. 1993; 114(3):510–532. [PubMed: 8272468]
- Rescorla RA, Heth CD. Reinstatement of fear to an extinguished conditioned stimulus. *Journal of Experimental Psychology: Animal Behavior Processes*. 1975; 1(1):88–96. [PubMed: 1151290]
- Sakamoto Y, Jones M, Love BC. Putting the psychology back into psychological models : Mechanistic versus rational approaches. *Memory & Cognition*. 2008; 36(6):1057–1065. [PubMed: 18927024]
- Sanborn AN, Griffiths TL, Navarro DJ. Rational approximations to rational models: Alternative algorithms for category learning. *Psychol Rev*. 2010; 117(4):1144–67. [PubMed: 21038975]
- Sissons HT, Miller RR. Spontaneous recovery of excitation and inhibition. *Journal of Experimental Psychology: Animal Behavior Processes*. 2009; 35(3):419–26. [PubMed: 19594286]
- Speekenbrink M, Shanks DR. Learning in a changing environment. *Journal of Experimental Psychology: General*. 2010; 139(2):266–98. [PubMed: 20438252]
- Summerfield C, Behrens TE, Koechlin E. Perceptual classification in a rapidly changing environment. *Neuron*. 2011; 71(4):725–36. [PubMed: 21867887]
- Tversky A, Kahneman D. Judgment under Uncertainty: Heuristics and Biases. *Science*. 1974; 185(4157):1124–31. [PubMed: 17835457]
- Wasserman, Ea; Castro, L. Surprise and change: variations in the strength of present and absent cues in causal learning. *Learning & Behavior*. 2005; 33(2):131–46. [PubMed: 16075834]
- Weiss DJ, Gerfen C, Mitchel AD. Speech Segmentation in a Simulated Bilingual Environment: A Challenge for Statistical Learning? *Language Learning and Development*. 2009; 5(1):30–49. [PubMed: 24729760]
- Wilson RC, Niv Y. Inferring relevance in a changing world. *Frontiers in Human Neuroscience*. 2012; 5(January):189. [PubMed: 22291631]
- Yildirim I, Degen J, Tanenhaus MK, Jaeger TF. Talker-specificity and adaptation in quantifier interpretation. *Journal of Memory and Language*. 2016; 87:128–143. [PubMed: 26858511]
- Yu AJ, Cohen JD. Sequential effects: Superstition or rational behavior? *Advances in Neural Information Processing Systems*. 2008; 21:1873–1880. [PubMed: 26412953]
- Yu AJ, Dayan P. Uncertainty, neuromodulation, and attention. *Neuron*. 2005; 46(4):681–92. [PubMed: 15944135]
- Zinszer, BD.; Weiss, DJ. When to Hold and When to Fold: Detecting Structural Changes in Statistical Learning. In: Knauff, M.; Pauen, M.; Sebanz, N.; Wachsmuth, I., editors. *Proceedings of the 35th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society; 2013. p. 3858–3863.

Appendix Implementation of the Ideal Learner

We implemented the ideal learner using the “Hibachi Grill Process” (HGP) mixture model (see Qian & Aslin, 2014 for details of the HGP model). The HGP mixture model is a Bayesian nonparametric model based on a three-step generative process: First, the HGP mixture model determines the length (i.e., the number of trials) of an upcoming bundle by drawing a sample from an arbitrary distribution. To match the experimental procedure used to generate mole sequences in the Whack-The-Mole game, we choose the Poisson

distribution. The mean of a Poisson distribution λ can be intuitively understood as the expected length l_i of a stimulus bundle i :

$$l_i \sim \text{Poisson}(\lambda) \quad (5)$$

Then in the second step, we sample the identity of this stimulus bundle from a Chinese Restaurant Process (CRP) with concentration parameters α :

$$p(\text{bundle identity}=k) \propto \begin{cases} n_k & \text{if identity } k \text{ exists} \\ \alpha & \text{if identity } k \text{ is novel} \end{cases} \quad (6)$$

where n_k is the number of bundles that have already been assigned with identity k . To match the experimental procedure, we set α to 2 in the ideal learner as well. These first two steps constitutes the HGP prior, which assigns a class identity to a sequence of observations simultaneously (i.e., the identity of an entire bundle), rather than to each observation individually as in the CRP prior.

The third and final step is to sample the actual l_i stimuli $\mathcal{T}_{ik}$ of bundle i , which has identity k , in one single pass. The distribution of stimuli is modeled by a categorical distribution with parameters θ_k :

$$\mathcal{Y}_{ik} \sim \text{Categorical}(\theta_k) \quad (7)$$

Since the Whack-The-Mole experiment was originally designed specifically for the HGP mixture model, we note that Equations (5) (6) (7) implement a generative process that is identical to the experimental procedure described above.

Inference via incremental iterative stochastic search

Given a sequence of mole locations as input and the above generative process as the assumption, the ideal learner infers bundle boundaries, bundle identities, and the parameters of the associated categorical distributions. As in the case of the idealistic learners, we aim to create an incremental ideal learner that predicts the reaction time at each trial for each subject. Due to the fact that each Whack-The-Mole session typically ran for a total of at least 1,000 trials, producing a prediction on each trial is a computationally intensive task and extremely challenging. The Gibbs sampling approach used in Qian and Aslin (2014) would be impractical as the time to wait for posterior samples to converge would scale superlinearly as a function of trial length.

We instead adopted a heuristic stochastic search procedure based on Gibbs sampling. After observing the mole's location at trial n , the ideal learner starts the inference process using the Gibbs sampling algorithm on the sequence of locations from Trial 1 to Trial n . During this process, we only record the posterior sample if it represents an improvement in the joint loglikelihood of data and the log-prior probability of the model over the last recorded best

sample (thereby controlling for both model fit and model complexity simultaneously). The search algorithm continues until no improvement is achieved for 200 iterations, with a maximum of 2,000 iterations. The best posterior sample given a sequence of mole locations up to the n th trial is then used as the representation of bundles to generate a prediction for the RT at the $(n + 1)$ th trial. This search is an approximation of the ‘ideal’ representation of bundles at each trial due to its stochastic nature. Compared to a particle-filter based algorithm, we have discovered that it is much less susceptible to being stuck on the local maxima while sacrificing slightly more computation time. This stochastic search algorithm is applied to the stimulus sequence of each subject in both experiments.

Figure 1.

A screenshot of the Whack-the-Mole game used in our experiments. On each trial, the mole appears at one of the four locations, and the positions of the background animals are randomly reshuffled.

Figure 2.

A schematic illustration of the underlying generative process of the Whack-The-Mole game. The entire sequence of stimuli in an experimental session consists of multiple stimulus bundles, each corresponding to a statistical pattern of some parameter settings (color-coded). Neighboring stimulus bundles always implement different patterns, and two different bundles may share the same pattern. Details on the length of a stimulus bundle and how statistical patterns reoccur (for example, the “green” pattern in this figure) are introduced in the method section.

Figure 3. In Experiment 1, for a given participant, each unique proportional distribution of the eight animals in the background was paired with an arbitrary statistical pattern. The green distribution 55-25-5-15% in this example was uniquely associated with the 4-1-2-1 (dinosaur-rabbit-hippopotamus-snail) animal configuration throughout the course of an experiment.

Figure 4.

Using the trial history of one typical subject as an example, the global learner (left) and the local learner (right) make different predictions regarding the probability of the mole's location on each trial. Each curve represents the probability that the mole is predicted to appear at one of the four holes in the game.

Figure 5.

Log RT vs. predictions of the idealistic learners (A and B) and the ideal learner (C) in 4 sample participants. Red dots represent data points excluded from analysis. Blue line represents the linear correlation between log RT and predictions of each model.

Figure 6.

On average, human behavior is described by the local learner and the ideal learner better than the global learner in conditions both with and without correlated visual cues. A higher R^2 value indicates a better fit, while a lower Kolmogorov-Smirnov (KS) value indicates a better fit. Error bars indicate between-model (and within-subjects) standard errors.

Figure 7.

Over the course of an experimental session (cutoff at 25 bundles $\sim 1,000$ trials), learners behaved increasingly more like the local learner (and the ideal learner) than the global learner with and without correlated visual cues (a higher R^2 or a lower Kolmogorov-Smirnov statistic). Model fits *between experiments* are not comparable.

Figure 8.

Scatterplot of human learners' reaction times as a function of model predictions on bundle-starting trials, using the first 5 trials of all bundles in Experiment 1 as an example. The predictions of the ideal learner provide a much more accurate description of the gradient of human RTs than the local learner.

Figure 9.

Within a stimulus bundle (defined with respect to true boundaries unknown to human learners), the ideal learner described RTs much better than the local learner for bundle-starting trials, especially when assessed by the Kolmogorov-Smirnov statistic. Model fits *between experiments* are not comparable.

Figure 10.

Over the 10 recurring appearances of an example probability distribution (nth stimulus is defined w.r.t. exact boundaries unknown to human learners), the local learner estimates the parameters from scratch each time, resulting in more volatile patterns in the most probable (45% of the time) and the second most probable locations (36% of the time). In comparison, predictions of the ideal learner quickly converged onto estimates that are much closer to the actual distribution during later recurrences, especially for the second most probable location. Panels C and D show that the predictions of the ideal learner correlate with RTs negatively ($\beta = -0.29$, $p < 0.001$), but those of the local learner do not ($\beta = -0.01$, $p > 0.5$).

Figure 11.

When contextual visual cues are present (Experiment 1 only), the ideal learner predicted human RTs better than the local learner in stimulus bundles with a less salient cue (high-entropy condition on the left), whereas the local learner did better in stimulus bundles with a more salient cue (low-entropy condition on the right).