



# Emotional Faces Facilitate Statistical Learning

Rista C. Plate<sup>1</sup> · Anna C. Schapiro<sup>1</sup> · Rebecca Waller<sup>1</sup>

Received: 14 September 2021 / Accepted: 7 June 2022  
© The Society for Affective Science 2022

## Abstract

Detecting regularities and extracting patterns is a vital skill to organize complex information in our environments. Statistical learning, a process where we detect regularities by attending to relationships between cues in our environment, contributes to knowledge acquisition across myriad domains. However, less is known about how emotional cues—specifically facial configurations of emotion—influence statistical learning. Here, we tested two pre-registered aims to advance knowledge about emotional signals and statistical learning: (1) we examined statistical learning in the context of emotional compared to non-emotional information, and (2) we assessed how emotional congruency (i.e., whether facial stimuli conveyed the same, or different emotions) influenced regularity extraction. We demonstrated statistical learning in the context of emotional signals. Further, we showed that statistical learning occurs more efficiently in the context of emotional faces. We also established that congruent cues benefited an online measure of statistical learning, but had varied effects when statistical learning was assessed via post-exposure recognition test. The results shed light on how affective signals influence well-studied cognitive skills and address a knowledge gap about how cue congruency impacts statistical learning, including how emotional cues might guide predictions in our social world.

**Keywords** Emotion · Statistical learning · Congruency · Facial expressions

Detecting regularities and extracting patterns is critical for making sense of complex environments. This skill, referred to as statistical learning, helps us to track, organize, and make predictions about auditory, visual, and tactile cues (see Frost et al., 2015; Krogh et al., 2013; Santolin & Saffran, 2018; Siegelman et al., 2017 for reviews). Specifically, the ability to notice sequential patterns of co-occurrence (i.e., things that occur together) helps to guide attention, form representations, make inferences, and predict future events (Sherman et al., 2020). For example, after participants viewed sequential triplets of emotion-evoking images, they were better at recognizing negative (e.g., soldier with gun) and positive (e.g., cute baby) image sequences than neutral image sequences (e.g., boy playing chess; Everaert et al., 2020). However, *evoking* emotion is only one route through which social information could promote statistical learning. No prior studies have

explored statistical learning in relation to one of the most frequently utilized social inputs that adults encounter—emotional facial signals (Barrett et al., 2019; Zebrowitz, 2017). Prior work has emphasized “supervised” learning of emotions that requires explicit instruction (e.g., language or labeling; Hoemann et al., 2019). However, facial emotion cues are incredibly complex (i.e., over 40 muscles in the human face), vary based on anatomy and expressivity style of an individual, and can be influenced by cultural or situational norms. Synthesizing, detecting signal, and making predictions from these stimuli present a major challenge to social agents. Thus, perceivers may need to rely on *unsupervised* learning mechanisms suited for contexts in which there are many and varied stimuli.

Beginning in infancy, faces dominate early visual input (Smith et al., 2018), give cues about the environment (Mumme et al., 1996; Sorce et al., 1985), and provide information about how to interact with the physical world (Moses et al., 2001; Patzwald et al., 2018). Facial cues impact learning, at least partly, by directly evoking an emotional response (e.g., fear to avoid environmental threats; Sorce et al., 1985; Everaert et al., 2020). However, learners could use “emotion as information” in the absence of having an emotional response evoked. That is, emotion cues could also provide

---

Handling Editor: Jonathan Gratch

✉ Rebecca Waller  
rwaller@sas.upenn.edu

<sup>1</sup> Department of Psychology, University of Pennsylvania, Levin Building, 425 S. University Ave, Philadelphia, PA 19104, USA

epistemic signals to guide learning (Wu et al., 2021). Thus, studies are needed to test the hypothesis that facial cues of emotion represent a mechanism that promotes statistical learning in social contexts, which could contribute to a growing body of research that has explored how social cues influence learning and inferences made through repeated social interactions (Gweon, 2021; Wu et al., 2021), including reward learning (Plate et al., 2021), exploration (Wu & Gweon, 2021), and curiosity (Dubey et al., 2021). One prior study investigated sequential statistical learning of emotional face signals, finding that infants track associations between emotive cues and subsequent behavior (social partners looking towards each other; Mermier et al., 2022). This evidence implicates an early role for statistical learning in the domain of interpersonal emotion communication.

At the same time, statistical learning in socioemotional contexts is complicated because learners have to discern regularities from an array of perceptual features and the relevant features and regularities among those features are often unlabeled (Ruba & Repacholi, 2020). Many features influence what learners attend to, but little research has investigated whether regularities are tracked in the emotion domain (see Ruba et al., 2022). Further, although people typically exhibit emotional responses consistent with the communicative context (Gergely & Király, 2019), there can be variability in expressivity within a single learning environment (e.g., classroom or office). That is, people can provide different emotional signals because of individual differences in expressivity (Niedenthal et al., 2017) or to “hide” or convey an expression that differs from their internal state (Ekman & Friesen, 1975). In terms of statistical learning, the presence of emotion cues that conflict with the social context could impede learning. Alternatively, incongruent emotional cues could evoke surprise and signal need for greater attention to evaluate potentially unseen information (Wu et al., 2017). These competing predictions highlight the need for studies that explore how learners interpret and learn from incongruent emotion cues (Wu et al., 2021). Outside the domain of facial expressions of emotion, incongruency impedes statistical learning (e.g., when audio and visual streams of triplets misalign; Glicksohn & Cohen, 2013), whereas congruent cues enhance statistical learning (e.g., when audio and visual streams of triplets align; Glicksohn & Cohen, 2013; Mitchel & Weiss, 2014). However, no prior research has investigated the effects of emotional congruency conveyed through facial expressions of emotion on statistical learning.

In the current study, we aimed to better understand how emotion signals, specifically facial configurations of emotion, impact statistical learning. Our task featured images of emotional faces (happy, angry, sad, fearful) organized into triplets, following classic statistical learning designs (e.g., Everaert et al., 2020; Saffran et al., 1996). Leveraging a well-established paradigm allowed us to contextualize findings

within the statistical learning literature more broadly. We assessed online statistical learning via reaction times to clicking on the second and third images of the triplet relative to the first image (Siegelman et al., 2018). We also quantified statistical learning using a recognition measure to assess whether participants identified triplets they had viewed in an exposure stream relative to novel, “foil” triplets at the end of the experiment. We hypothesized that participants would evidence statistical learning in the context of emotional faces, as shown through their reaction times and on the recognition test. Second, we explored whether emotion cues (versus social information more broadly) affect statistical learning, by including a comparison condition where participants viewed triplets of neutral facial expressions. We hypothesized that statistical learning would be enhanced in the context of emotional faces relative to the comparison condition. Third, we investigated the impact of congruency on statistical learning in the context of emotional faces by including triplets in which models either depicted the same emotion (congruent) or two different emotions (incongruent). Accordingly, we include a comparison condition where participants were shown triplets of abstract shapes that also included congruent (all one color) or incongruent (two colors) information within triplets. We hypothesized that congruent information would facilitate statistical learning and that incongruent emotional information would impair statistical learning.

## Method

The study was preregistered at: <https://aspredicted.org/mp9du.pdf>. The experiment, de-identified dataset, and analysis script are on Open Science Framework (<https://osf.io/gc3u5/>; see Supplemental Materials for information regarding “Analysis Tools”).

## Participants

Participants were 327 adults (191 female, 131 male, 1 selected “other” to describe gender, 3 indicated “prefer not to say”;  $M_{age} = 20.00$  years,  $SD_{age} = 1.76$  years; 89 Asian, 25 Black or African American, 24 Hispanic or Latinx, 21 Multiracial, 167 White) who were recruited from an undergraduate subject pool at a large university in the Northeast region of the USA. All participants provided informed consent and received course credit for their participation. The Institutional Review Board approved the research.

## Design

Participants were randomly assigned to complete one of the three task conditions: *emotional face* ( $n=108$ ; 66 female, 41

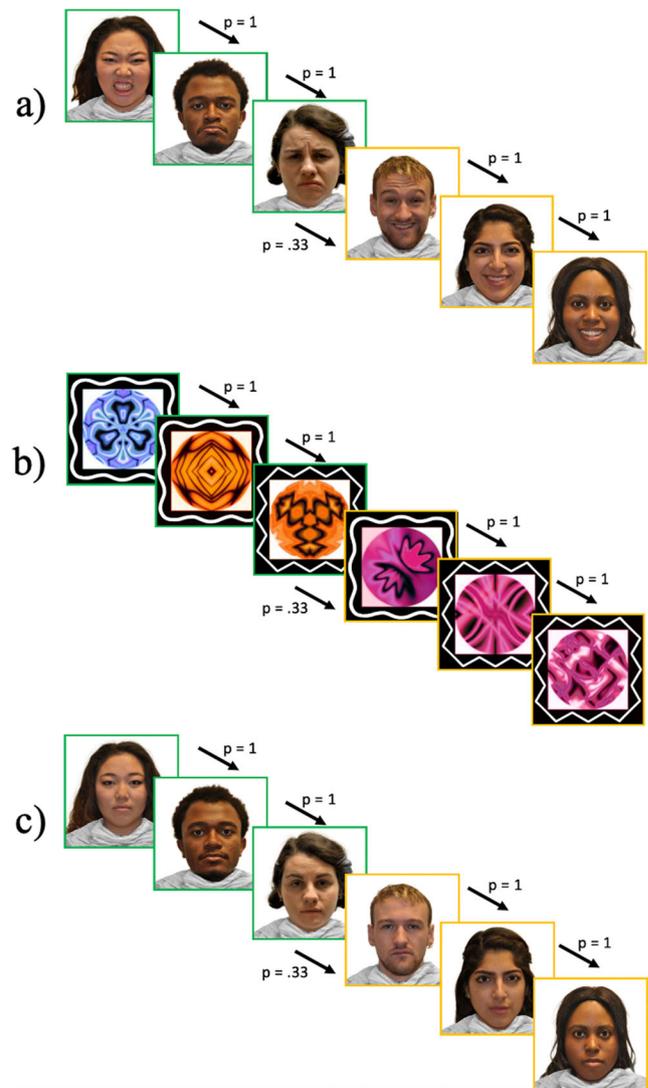
male, 1 indicated “prefer not to say”;  $M_{age} = 19.78$  years,  $SD_{age} = 1.33$  years; 31 Asian, 9 Black or African American, 8 Hispanic or Latinx, 6 Multiracial, 54 White), *neutral face* ( $n=108$ ; 66 female, 40 male, 1 indicated “prefer not to say”;  $M_{age} = 20.00$  years,  $SD_{age} = 2.10$  years; 25 Asian, 12 Black or African American, 9 Hispanic or Latinx, 8 Multiracial, 53 White), or *shape* ( $n=111$ ; 59 female, 50 male, 1 selected “other” to describe gender, 1 indicated “prefer not to say”;  $M_{age} = 20.22$  years,  $SD_{age} = 1.75$  years; 33 Asian, 4 Black or African American, 7 Hispanic or Latinx, 7 Multiracial, 60 White). Participant age ( $p = .18$ ), gender ( $p = .36$ ), and race ( $p = .76$ ) did not differ by condition. The G\*Power calculation for three predictors (condition, congruence, and time) assuming a small effect (small effect size chosen based on Siegelman et al., 2018 reporting an  $R^2$  of 0.23 for their online statistical learning measure) recommended a sample size of 77 for each condition, which we exceeded (i.e., to allow for the possibility of exclusions).

**Triplet Design** Stimuli (i.e., individual images of faces or shapes, see below) in the emotional face and shape conditions were randomly assigned to an emotion/color (Figure 1 A and B). These stimuli were also randomly assigned to a triplet with the constraint that two triplets be congruent (i.e., each image in the triplet has the same emotion/color) and two triplets be incongruent (i.e., defined as the first image in the triplet being of a different emotion/color than the second and third images). (Note: within valence emotions were allowed within incongruent triplets and we report the results of analyzing the effects of valence in the Supplemental Materials.) In the neutral face condition, the stimuli were randomly assigned to triplets (Figure 1C). Triplets were then randomized to create a continuous stream with the constraint that the same triplet could not appear twice in a row across 48 triplet presentations.

## Stimuli and Materials

**Face Stimuli** We selected twelve models each producing four facial configurations (happy, angry, fearful, and sad) from RADIATE, a validated stimulus set that includes a diverse representation of models (Conley et al., 2018; Tottenham et al., 2009). Models in the stimulus set were provided instructions for how to create each facial configuration and were given time to practice their expression using a mirror (see Conley et al., 2018 for details). We included four White, four Black, two Asian, and two Hispanic models that were balanced for gender. The neutral face from each model was included for the neutral face condition.

We purposefully selected images for emotions that reliably belonged to a particular category of emotion and that were distinct from each other. Thus, we could ensure that stimuli in the emotional face condition were separable from those in the neutral face condition *and* that the emotions were distinguishable within incongruent triplets. Our selection of models



**Fig. 1** Example triplet configurations for the emotional face (A), shape (B), and neutral face (C) conditions. Note: Relations within triplets are  $p = 1$  and relations between triplets are  $p = .33$ . Triplets are outlined in green and yellow here for illustration purposes only; there were no cues—other than the probabilistic relations—to the triplets for participants. Face stimuli are from the RADIATE stimulus set (Conley et al., 2018; Tottenham et al., 2009)

was narrowed using validation ratings from RADIATE set to include those with Kappa  $>.59$  for at least one expression of happy, sad, fear, and angry (open or closed mouth). Convergence across stimuli indicated that open-mouthed happy (Kappa range=.74–.98), angry (Kappa range=.62–.97), and fear (Kappa range=.60–.85) and closed-mouthed sad (Kappa range=.60–.90) faces had the best reliability. When comparing models that fit criteria, our final selection of models was those with the highest average proportion correct identification rating: AF07 (.89), AM04 (.83), BF06 (.74), BF10 (.73), BM03 (.82), BM04 (.88), HF02 (.94), HM09 (.76), WF02 (.89), WF14 (.85), WM02 (.81), and WM04 (.82).

**Shape Stimuli** Shape stimuli were adapted from prior research (Schapiro et al., 2012). Twelve unique stimuli were edited in Procreate to create four versions delineated by color (blue, pink, yellow, and orange). There were two border designs added to the stimuli (characterized by a smooth line and a jagged line). The border varied across stimuli but was not systematically manipulated within the task (i.e., similar to model gender in face conditions). Although facial and non-facial stimuli cannot be fully equated, the images were complex designs that included many features that learners could attempt to track in addition to color and specific image (e.g., border, design elements), thus maximizing their utility use as a contrast condition. As in daily living environments, there are myriad features that learners could attend to, and we attempted to include stimuli (in all conditions) with a number of these features in order for participants to have to glean the regularities over time.

## Procedure

Participants completed the task online. In an initial exposure phase, participants viewed one stimulus at a time. They were instructed to notice any co-occurrences and told that they would be tested at the end (see Arciuli et al., 2014 and Siegelman & Frost, 2015 for discussion of explicit instructions in statistical learning paradigms. We included these instructions to maintain consistency with Siegelman et al., 2018). Participants clicked on each image to proceed to the next, advancing the task at their own pace (participants saw 144 images in total; i.e., 48 triplet presentations).

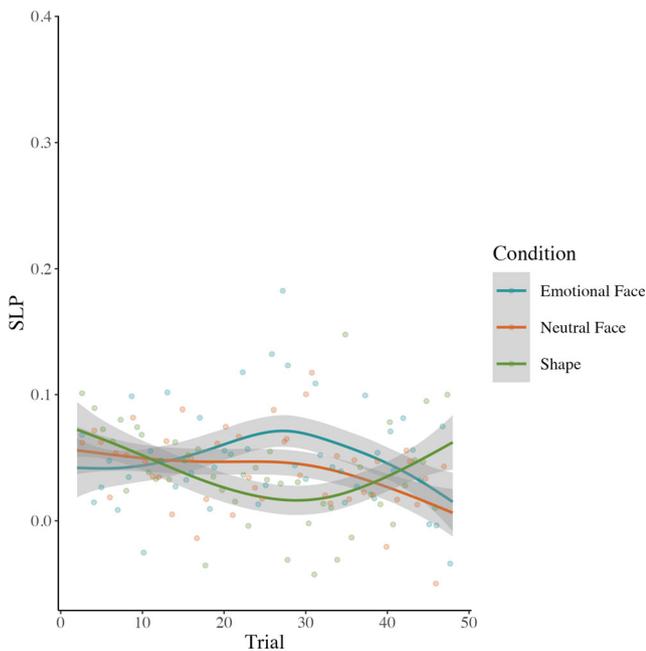
Next, participants completed a testing phase in which they viewed streams of three images and were asked whether the streams were familiar or novel. This phase included three types of triplets: “Target triplets” were the exact four triplets viewed during the exposure phase. “Foil triplets” included the images from triplets presented in the exposure phase, but in a novel order. Specifically, images in foil triplets had the same position to that from the exposure phase, but were combined with two new images (e.g., if exposure phase triplets were ABC, DEF, and GHI, then a test foil triplet could be AEI). “Part-triplets” had the same images in positions 2 and 3 as those presented in the exposure phase but were combined with a different position 1 image (e.g., a part-triplet from the example above might be AHI). There were four target triplets presented during the exposure phase. Thus, we included four foil triplets and four part-triplets during the recognition testing phase. Each triplet was presented three times (in a random order) for a total of 36 test trials. The dependent variable was whether the participant identified the triplet as familiar or novel. Target triplets were familiar to the participant (i.e., they had been seen during the exposure phase) and foil triplets and part-triplets were novel.

## Analyses

**Preprocessing** Six participants were excluded based on reaction time (RT) criteria ( $\pm 3$  *SD* of sample mean; preregistered). Trials that exceeded  $\pm 2$  *SD* from individual participant means were trimmed to the cutoff value (3% of trials; preregistered). We used an online measure of “Statistical Learning Performance” based on prior research (SLP; Siegelman et al., 2018). Statistical Learning Performance was calculated by subtracting the mean log RT of stimuli in the second and third positions of triplets from the log RT of the stimulus in the first position of the triplets. For example, for a given triplet, ABC, Statistical Learning Performance would be calculated as  $\log RT(A) - \log RT(\text{mean}(B,C))$ . We log-transformed RT and used a measure that calculated RT differences within participants to reduce concerns about device-related RT effects in our online sample (as discussed in Siegelman et al., 2018). Accordingly, we obtained a measure of Statistical Learning Performance for each of the 48 triplet presentations. Scores greater than 0 reflect learning of the statistical patterns via a relative increase of reaction time speed within, as compared to between, triplets. Calculating Statistical Learning Performance for each triplet presentation allowed us to examine statistical learning across trials.

**Deviations from Preregistration** Upon visualization of the data, we observed two patterns that required changes to our preregistered analyses. First, to make claims about differences in *learning*, we included time (i.e., triplet trial number) as a predictor in all analyses of the exposure phase. Including time was specified for some, but not all, analyses in the preregistration (note: a lag followed participants having to click to start the experiment; thus, we excluded trial 1 from analyses to reduce any artificial inflation of Statistical Learning Performance). Second, the patterns across time were not linear (Figure 2). Specifically, the shapes of the curves across conditions and time differed, with a peak at trial 28. To address this issue, we examined the data in two halves: an initial period (trials 2–28, “initial learning period”) and a second experimental period (trials 29–48, “later learning period”). We concentrate our presentation of the results based on dividing the data in this way. However, the results from all preregistered analyses are available in the Supplemental Materials (“Complete Preregistered Analyses”) and are largely consistent with the results presented in the manuscript.

**Models Assessed** To test whether participants showed evidence of statistical learning in the context of emotional faces during the exposure phase, we regressed Statistical Learning Performance on time for the initial and later learning periods. Next, we examined whether the congruence of the triplets influenced statistical learning in the emotional face and shape conditions (because the neutral face condition only included



**Fig. 2** Statistical Learning Performance by trial and condition. Note: Error bars are standard error and points are averages by triplet trial number and condition

congruent trials). For the initial and later learning periods, we regressed Statistical Learning Performance on the interaction between condition (shape=-.5, emotion=.5), congruence (incongruent=-.5, congruent=.5), and time (i.e., trial number, mean-centered), and all lower order effects (i.e., main effects of condition, congruence, time, and the two-way interactions). We included a by-participant random intercept and by-participant random slopes for congruence and time. We further investigated the effect of congruency to better understand whether congruent information improved statistical learning. We compared Statistical Learning Performance for congruent triplets in the emotional face and shape conditions versus the neutral face condition, which did not have contrasting congruent and incongruent triplet information. For each the initial and later learning phases, we regressed Statistical Learning Performance for congruent triplets on condition (emotional face as referent), time (mean-centered), and their interaction with a by-participant random slope for time.

In addition to online statistical learning, we used a recognition test at the end of the experiment. Using a logistic mixed-effects model to take advantage of multiple test trials, we regressed participant responses (“familiar”=1, “not familiar”=0) on condition (emotional face condition as referent), trial type (target, part-triplet, foil; part-triplet as referent), and their interaction. We included a by-participant random intercept and a by-participant random slope for trial type. We also conducted a one-way ANOVA to assess accuracy (defined as indicating that target triplets were familiar and foil triplets and part-triplets were novel) across conditions (pairwise contrasts

were Bonferroni-adjusted to account for multiple comparisons).

## Results

### Participants Show Online Evidence of Statistical Learning with Emotional Faces

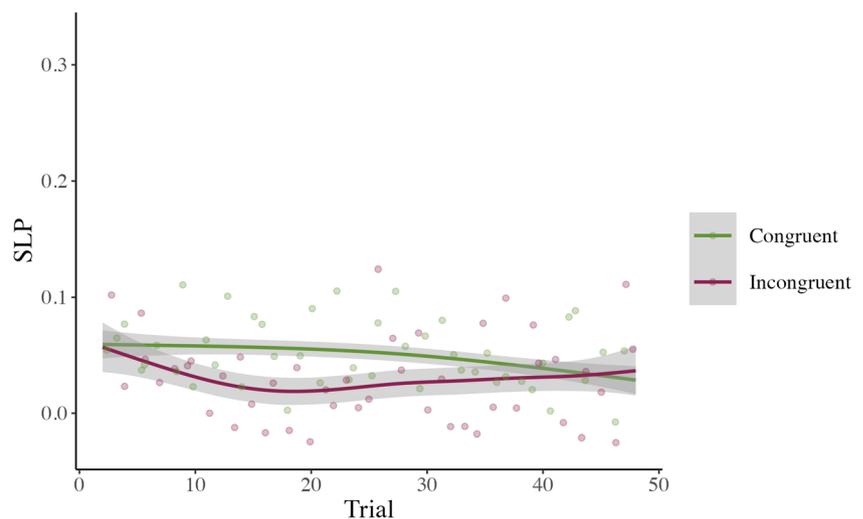
First, we examined whether there was evidence of statistical learning in the emotional face condition. In the initial learning period, Statistical Learning Performance increased over time for the emotional face condition,  $b = .002$ ,  $X^2(1) = 20.35$ ,  $p < .001$ . For completeness, we also examined the comparison conditions: Statistical Learning Performance did not change significantly over time for the neutral face condition ( $b = -.0003$ ,  $t = -0.65$ ,  $p = .519$ ) and decreased for the shape condition ( $b = -.002$ ,  $t = -4.41$ ,  $p < .001$ ). The relationship between Statistical Learning Performance and time also differed significantly across conditions,  $X^2(2) = 39.84$ ,  $p < .001$ ; Figure 2). Namely, in the emotional face condition, Statistical Learning Performance increased more during the initial learning period compared to the shape and neutral face conditions, shape vs. emotional face:  $b = -.003$ ,  $t = -3.63$ ; neutral face vs. emotional face:  $b = -.005$ ,  $t = -6.29$ ; main effect of condition was not significant,  $X^2(2) = 1.78$ ,  $p = .411$ ). Thus, statistical learning for the initial learning period was greater in the context of emotional faces compared to neutral faces and shapes.

In the later learning period, the relationship Statistical Learning Performance decreased for the emotional face condition,  $b = -0.002$ ,  $X^2(1) = 3.93$ ,  $p = .047$ , and the neutral face condition,  $b = -.003$ ,  $X^2(1) = 12.80$ ,  $p < .001$ , but increased in the shape condition,  $b = .003$ ,  $X^2(1) = 14.09$ ,  $p < .001$ . In line with these patterns, the interaction between time and condition was significant,  $X^2(2) = 28.73$ ,  $p < .001$ ). Statistical Learning Performance increased more over time in the shape condition compared to the emotional face,  $b = .005$ ,  $t = 4.01$ , condition, neutral face vs. emotional face:  $b = -.001$ ,  $t = -1.05$ ). Thus, we found evidence for statistical learning of shapes, but not emotion or neutral faces, for the later learning period of the experiment.

### Statistical Learning Is Better for Congruent Versus Incongruent Stimuli

Next, we examined whether congruency influenced statistical learning in the emotional face and shape conditions. There was a main effect of congruence, such that Statistical Learning Performance was higher for congruent triplets;  $b = -0.04$ ,  $X^2(1) = 8.69$ ,  $p = .003$ . The main effect of congruence was qualified by an interaction with time for the initial learning period,  $b = 0.001$ ,  $X^2(1) = 5.21$ ,  $p = .022$ ; Figure 3, such that Statistical Learning Performance decreased more over

**Fig. 3** Statistical Learning Performance by trial and congruence. Note: Error bars are standard error and points are averages by triplet trial number and condition



time for incongruent,  $b = .0009$ ,  $X^2(1) = 3.13$ ,  $p = .08$ , versus congruent,  $b = .0008$ ,  $X^2(1) = 2.48$ ,  $p = .12$  triplets. The interaction was not significant for later learning trials,  $b = -0.001$ ,  $X^2(1) = 0.49$ ,  $p = .483$ . The three-way interaction between condition, congruence, and time was not significant for either learning period (Table 1).

### Congruent Emotional Information Facilitates Statistical Learning

To further understand the role of congruence, we examined congruent trials only, comparing all three conditions (i.e., neutral face condition was included here since it only contained congruent triplets). In the initial learning period, the interaction between condition and time was significant,  $X^2(2) = 10.46$ ,  $p = .005$ . Participants in the emotional face condition showed a relative increase in Statistical Learning Performance for congruent triplets compared to those in the shape and neutral face conditions (shape vs. emotional face:  $b = -.006$ ,  $t = -3.16$ ; neutral face vs. emotional face:  $b = -.003$ ,  $t = -2.25$ ), reflecting the overall pattern described above. There was no

difference between the shape and neutral face conditions ( $b = -.002$ ,  $t = -1.04$ ), suggesting that the advantage was specific to congruent emotional faces.

The interaction between condition and time was also significant in later learning trials,  $X^2(2) = 7.58$ ,  $p = .02$ . There was a relative decrease in Statistical Learning Performance for congruent triplets in the emotion relative to the shape condition (shape vs. emotional face:  $b = .008$ ,  $t = 2.37$ ). However, there was no difference between the emotion and neutral face conditions ( $b = .0002$ ,  $t = 0.07$ ) and the shape condition increased compared to the neutral face condition ( $b = .007$ ,  $t = 2.43$ ). Together, these patterns suggest a protracted learning trajectory for shapes compared to emotional faces (Supplemental Materials Table S2 and Figure S1).

### Emotional Faces Facilitate Enhanced Specificity in Identifying Triplets at Test

Finally, we examined whether emotional faces influenced the ability to distinguish target triplets as familiar compared with part-triplets and foils. There was a main effect of triplet type,

**Table 1** Full model output for effect of condition (shape =  $-.5$ , emotional face =  $.5$ ), congruence (incongruent =  $-.5$ , congruent =  $.5$ ), and time (trial number, mean centered) on Statistical Learning Performance

Predictors	Initial learning period				Later learning period			
	<i>b</i>	<i>CI</i>	<i>t</i>	<i>p</i>	<i>b</i>	<i>CI</i>	<i>t</i>	<i>p</i>
(Intercept)	0.05	0.04 to 0.06	8.6	<0.001	0.04	0.02 to 0.05	4.25	<0.001
Condition	0.01	-0.01 to 0.04	1.27	0.205	0.01	-0.03 to 0.04	0.5	0.617
Congruence	0.04	0.01 to 0.06	2.95	0.003	0.02	-0.01 to 0.05	1.32	0.188
Time	0	-0.00 to 0.00	-0.11	0.913	0	-0.00 to 0.00	0.39	0.698
Condition-by-Congruence	0.03	-0.02 to 0.08	1.1	0.273	0.06	-0.00 to 0.12	1.85	0.064
Condition-by-Time	0	0.00 to 0.01	3.05	0.002	-0.01	-0.01 to -0.00	-2.25	0.024
Congruence-by-Time	0	0.00 to 0.00	2.28	0.022	0	-0.00 to 0.00	-0.7	0.483
Condition-by-Congruence-by-Time	0	-0.00 to 0.00	0.82	0.412	0	-0.01 to 0.00	-1.52	0.129

$X^2(2) = 161.20, p < .001$ . Across all conditions, participants were more likely to identify target triplets as familiar compared to part-triplets (target vs. part-triplet  $b = 2.22, z = 11.38, OR = 9.23$ ; Figure 4) and were more likely to identify part-triplets as familiar compared to foil triplets ( $b = -1.33, z = -8.26, OR = 0.26$ ), providing evidence for statistical learning. The main effect of condition was not significant,  $X^2(2) = 5.46, p = .065$ . However, there was an interaction between trial type and condition,  $X^2(4) = 19.80, p < .001$ . Participants were more likely to identify target triplets as familiar compared to part-triplets in the emotional face compared to the shape condition ( $b = -0.87, z = -3.22, OR = 0.42$ ). There was no difference between the neutral face and emotional face conditions ( $b = -0.40, z = -1.47, OR = 0.67$ ).

Results from a one-way ANOVA comparing accuracy across conditions indicated higher accuracy in the emotional face ( $Mean_{Accuracy} = .68$ ), as compared to shape ( $Mean_{Accuracy} = .63$ ) condition, omnibus:  $F(2, 324) = 3.07, p = .048$ ; emotion vs. shape:  $t(209.17) = 2.47, p_{adj} = .04$ . There were no differences in accuracy between the neutral face ( $Mean_{Accuracy} = .65$ ) condition and the emotional face or shape conditions (neutral face vs. emotional face  $p_{adj} = .47$ , neutral face vs. shape  $p_{adj} = .95$ ). Together, results establish evidence for statistical learning of emotional faces using the recognition measure and show that emotional faces lead to enhanced differentiation between targets and part-triplets.

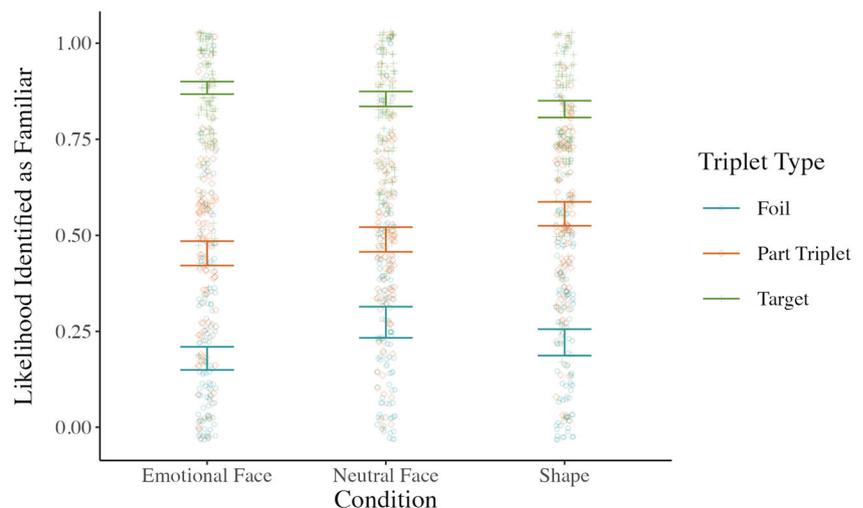
### Exploratory Follow-up Analyses

To better understand performance in the exposure versus test phases, we ran two exploratory follow-up analyses. First, we examined target triplets, part-triplets, and foils in the emotional face and shape conditions to test whether congruency influenced test performance. For target triplets, there was a main effect of congruence,  $b = 0.98, X^2(1) = 12.61, p < .001, OR = 2.68$ , with participants more likely to accurately identify

congruent targets as familiar. Effects of condition,  $X^2(1) = 2.86, p = .091$ , and the interaction between condition and congruence,  $X^2(1) = 0.0005, p = .982$  were not significant. For part-triplets, there were main effects of congruence and condition, interaction was not significant,  $X^2(1) = 0.03, p = .860$ . Here, the main effect of congruence suggested that participants were more likely to misidentify congruent part-triplets as familiar compared to incongruent triplets,  $b = 0.72, X^2(1) = 12.69, p < .001, OR = 2.06$ . Additionally, participants were more likely to misidentify part-triplets as familiar in the shape versus emotion condition,  $b = -0.62, X^2(1) = 6.75, p = .009, OR = 0.54$ . For foils (for which there were only incongruent triplets), there was a marginal effect of condition suggesting that participants were more likely to misidentify shape versus emotion foils as familiar,  $b = -1.08, X^2(1) = 3.30, p = .069, OR = 0.34$ .

Second, we tested whether Statistical Learning Performance during exposure related to test performance by regressing the likelihood of identifying targets as familiar on the interaction between condition, Statistical Learning Performance during exposure, and congruence, with a by-participant random intercept and by-participant random slope for congruence. The effect of Statistical Learning Performance was significant,  $b = 4.07, X^2(1) = 9.49, p = .002, OR = 58.73$ , suggesting that better online Statistical Learning Performance was associated with higher accuracy at test. The interaction between Statistical Learning Performance and congruence was also significant, indicating a stronger relationship between Statistical Learning Performance and accuracy for congruent triplets,  $b = 7.39, X^2(1) = 8.85, p = .003, OR = 1625.78$ ; Figure S2. There was no effect of condition ( $p > .20$ ; see Supplemental Materials Table S3 for full model output). To test whether the relationship between Statistical Learning Performance and accuracy at test held for all three conditions, we examined the simple correlation between Statistical Learning Performance and accuracy in the neutral face condition, which was significant ( $r = .22$ ,

**Fig. 4** Likelihood of saying that a triplet was familiar at test. Note: Points are individual averages, error bars are  $\pm$  one standard error. The effect of triplet type was significant with differences between target and part-triplets and part-triplets and foils. Additionally, participants were more likely to identify target triplets as familiar, as compared to part-triplets, when viewing emotional faces as compared to shapes



$p = .02$ ). Additionally, there was no interaction with condition in the model when all three conditions were included,  $\chi^2(2) = 1.59$ ,  $p = .45$ , indicating consistency of the relationship across conditions.

## Discussion

In a paradigm using emotional facial cues, we provide evidence for statistical learning. We extend prior literature (Everaert et al., 2020) and establish statistical learning of emotional faces using both an online measure of statistical learning and a post-exposure, recognition measure. We find evidence suggesting an advantage of emotional faces for tracking underlying patterns relative to neutral faces or shapes. Unlike prior research (i.e., Everaert et al., 2020), our stimuli were not intended to provoke strong or conscious emotional reactions in participants. Thus, our findings are consistent with recent evidence suggesting that individuals use “emotion as information” (Wu et al., 2021) to guide learning of their physical and social environments and that the distribution of facial signals can guide social inference (Dotsch et al., 2016).

Our results raise the possibility that different stimuli are related to different trajectories of learning: evidence of learning occurred earlier in the experiment for emotional faces whereas evidence of learning occurred later for shapes. Emotion cues may especially be useful early on, when learners attempt to resolve ambiguity or uncertainty in the learning environment (Clément & Dukes, 2017; Walle et al., 2017). This idea is consistent with the somewhat surprising finding that Statistical Learning Performance actually decreased for emotional faces in the latter half of the experiment, which may be explained by participants having already learned the associations. Together, results suggest that emotional signals enhance statistical learning, raising the possibility that the social context provides an environment that can uniquely facilitate one type of learning that is essential for understanding, and making inferences about, the world (Aslin, 2017). An exciting future direction is to explore individual differences, which have been shown both in terms of attention to statistical information (Hanson et al., 2017; Harms et al., 2018; Van de Cruys et al., 2013, 2014) and use of emotional cues (e.g., Chaplin & Cole, 2005).

In addition to the overall effects of emotional faces on statistical learning, we found that congruency benefited statistical learning. Congruency facilitated initial statistical learning, and higher statistical learning of congruent triplets was associated with higher accuracy at test. However, the effects of congruency at test were not entirely straightforward. While congruency improved participants’ ability to correctly identify target triplets as familiar, congruent part-triplets were more readily misidentified as familiar. Therefore, congruency may be associated with a familiarity bias more generally, and could

potentially mislead a learner to infer patterns that are inconsistent with patterns in the environment. More research is needed to understand the role of congruency of facial cues for statistical learning. For example, we defined congruency based only on the specific emotion conveyed, but congruency could also be defined in terms of perceptual features (e.g., open or closed mouth; congruence between face and body postures, Meeren et al., 2005) or cross modal cues (e.g., vocalizations; Dolan et al., 2001). The level at which congruency is defined has the potential to influence its impact on statistical learning, particularly whether incongruence leads to confusion or evokes greater attentional orientating in service of exploration (Wu et al., 2017).

Our results should be considered alongside several limitations. First, we used static visual images of posed facial configurations of emotion thereby limiting ecological validity and minimizing granularity and depth of emotion signals (Atias et al., 2019; Barrett et al., 2019; Schirmer & Adolphs, 2017). To maximize reliability of the emotions captured in our stimuli, we also mixed closed and open mouth facial displays. In doing so, we may have inadvertently introduced additional perceptual differences across stimuli that impacted results. Nevertheless, we still found evidence that even relatively pared down emotional signals impacted statistical learning, supporting the idea that emotion can be viewed as an important informational signal that is not solely reliant on a robust emotional reaction in the learner. Future studies are needed to explore how other emotional information (e.g., vocalizations) contribute to evaluations of congruency—as emotion cues can be congruent (happy face and happy vocalization) or incongruent (happy face and sad vocalization) both within and across individuals (see Wu et al., 2017). Related to this point, we focused on sequential cues to capture one aspect of emotion environments. Future research could explore simultaneous cues within facial displays (e.g., by varying facial action units consistently or inconsistently with prototypical emotion signals).

Second, by using faces, participants may have tracked patterns using a strategy of “naming” the models. We attempted to mitigate this concern by including a neutral face condition, and there was no difference between the emotional and neutral face conditions in terms of accuracy. Thus, it is possible that facial cues in general confer some advantage. This idea aligns with evidence in perceptual learning of faster learning of facial stimuli, which may be due to the familiarity or complexity of the stimuli (Fine & Jacobs, 2002). Exposure to specific individuals (Heron-Delaney et al., 2011; Tanaka et al., 2013) or emotions (Pollak et al., 2009) can also influence perceptual learning and affect visual expertise in perceptual categorization (Bukach et al., 2006; Tanaka et al., 2005). Thus, to the extent to which we observed similarities between the emotion and face conditions (i.e., in the test phase), we could look to perceptual expertise as a potential mechanism (Curby &

Gauthier, 2010; Tanaka & Taylor, 1991). Accounting for familiarity and complexity is one challenge when it comes to comparing facial and non-facial stimuli. Research may also need to account for perceptual differences (e.g., visual angle of the stimuli) in a more controlled testing environment that is not online.

Finally, we asked participants to attend to patterns in our instructions (“Some images tend to follow each other. Your task is to try and notice these co-occurrences”). While the instructions are consistent with previous research (Siegelman et al., 2018) and did not differ across conditions, we provided a task-related goal that may limit our knowledge of whether participants would have picked up on the patterns without any prior expectation that there were relationships between the stimuli.

In sum, we provide evidence of statistical learning in the context of emotional faces, highlighting a potential mechanism for how individuals discern patterns in a social world. These findings build on prior literature that has leveraged different domains (e.g., language, color sequences) to demonstrate that tracking patterns of co-occurrence influences the inferences we make to understand our complex environments. Here, we find that emotional faces may enhance our ability to notice such patterns, thereby supporting learning in socioemotional environments. Our findings inform how emotion signals may enhance in-the-moment learning and sharpen a domain general skill for use in social contexts, laying the foundation for understanding the complexity of how we adjust to our dynamic social world.

**Acknowledgements** We thank Natalie Corbett for help with data collection.

## Additional Information

**Funding Information** This research was supported by institutional funding from the University of Pennsylvania to R.W. A University of Pennsylvania MindCORE Postdoctoral Fellowship funded R.C.P.

**Data Availability** The experimental paradigm and de-identified data are available on Open Science Framework (<https://osf.io/gc3u5/>).

**Ethical Approval** The University of Pennsylvania Institutional Review Board approved the research.

**Conflicts of Interest** The authors have not conflicts of interest to declare.

**Informed Consent** All participants provided informed consent for participation.

**Code Availability** The analysis script is available on Open Science Framework (<https://osf.io/gc3u5/>).

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s42761-022-00130-9>.

**Author Contributions** All authors contributed to idea conception, task design, and preregistered analysis. RCP programmed the experimental task, collected data, performed the statistical analysis, and wrote the first draft of the manuscript. All authors were involved in manuscript revision.

## References

- Arciuli, J., von Torkildsen, J. K., Stevens, D. J., & Simpson, I. C. (2014). Statistical learning under incidental versus intentional conditions. *Frontiers in Psychology, 5*. <https://doi.org/10.3389/fpsyg.2014.00747>.
- Aslin, R. N. (2017). Statistical learning: A powerful mechanism that operates by mere exposure. *WIREs Cognitive Science, 8*(1–2), e1373. <https://doi.org/10.1002/wcs.1373>.
- Atias, D., Todorov, A., Liraz, S., Eidinger, A., Dror, I., Maymon, Y., & Aviezer, H. (2019). Loud and unclear: Intense real-life vocalizations during affective situations are perceptually ambiguous and contextually malleable. *Journal of Experimental Psychology: General, 148*(10), 1842–1848. <https://doi.org/10.1037/xge0000535>.
- Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest, 20*(1), 1–68. <https://doi.org/10.1177/1529100619832930>.
- Bukach, C. M., Gauthier, I., & Tarr, M. J. (2006). Beyond faces and modularity: The power of an expertise framework. *Trends in Cognitive Sciences, 10*(4), 159–166. <https://doi.org/10.1016/j.tics.2006.02.004>.
- Chaplin, T. M., & Cole, P. M. (2005). The role of emotion regulation in the development of psychopathology. In B. L. Hankin & J. R. Z. Abela (Eds.), *Development of psychopathology: A vulnerability-stress perspective* (pp. 49–74). Inc: Sage Publications. <https://doi.org/10.4135/9781452231655.n3>.
- Clément, F., & Dukes, D. (2017). Social appraisal and social referencing: Two components of affective social learning. *Emotion Review, 9*(3), 253–261. <https://doi.org/10.1177/1754073916661634>.
- Conley, M. I., Dellarco, D. V., Rubien-Thomas, E., Cohen, A. O., Cervera, A., Tottenham, N., & Casey, B. (2018). The racially diverse affective expression (RADIATE) face stimulus set. *Psychiatry Research, 270*, 1059–1067. <https://doi.org/10.1016/j.psychres.2018.04.066>.
- Curby, K. M., & Gauthier, I. (2010). To the trained eye: Perceptual expertise alters visual processing. *Topics in Cognitive Science, 2*(2), 189–201. <https://doi.org/10.1111/j.1756-8765.2009.01058.x>.
- Dolan, R. J., Morris, J. S., & de Gelder, B. (2001). Crossmodal binding of fear in voice and face. *Proceedings of the National Academy of Sciences, 98*(17), 10006–10010. <https://doi.org/10.1073/pnas.171288598>.
- Dotsch, R., Hassin, R. R., & Todorov, A. (2016). Statistical learning shapes face evaluation. *Nature Human Behaviour, 1*(1), 1–6. <https://doi.org/10.1038/s41562-016-0001>.
- Dubey, R., Mehta, H., & Lombrozo, T. (2021). Curiosity is contagious: A social influence intervention to induce curiosity. *Cognitive Science, 45*(2), e12937. <https://doi.org/10.1111/cogs.12937>.
- Ekman, P., & Friesen, W. V. (1975). *Unmasking the face: A guide to recognizing emotions from facial clues*. Prentice-Hall.
- Everaert, J., Koster, E. H. W., & Joormann, J. (2020). Finding patterns in emotional information: Enhanced sensitivity to statistical regularities within negative information. *Emotion, 20*(3), 426–435. <https://doi.org/10.1037/emo0000563>.

- Fine, I., & Jacobs, R. A. (2002). Comparing perceptual learning across tasks: A review. *Journal of Vision*, 2(2), 5–203. <https://doi.org/10.1167/2.2.5>.
- Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences*, 19(3), 117–125. <https://doi.org/10.1016/j.tics.2014.12.010>.
- Gergely, G., & Király, I. (2019). Natural pedagogy of social emotions. *Foundations of Affective Social Learning: Conceptualizing the Social Transmission of Value*, 87–114.
- Glicksohn, A., & Cohen, A. (2013). The role of cross-modal associations in statistical learning. *Psychonomic Bulletin & Review*, 20(6), 1161–1169. <https://doi.org/10.3758/s13423-013-0458-4>.
- Gweon, H. (2021). Inferential social learning: Cognitive foundations of human social learning and teaching. *Trends in Cognitive Sciences*, 25(10), 896–910. <https://doi.org/10.1016/j.tics.2021.07.008>.
- Hanson, J. L., van den Bos, W., Roeber, B. J., Rudolph, K. D., Davidson, R. J., & Pollak, S. D. (2017). Early adversity and learning: Implications for typical and atypical behavioral development. *Journal of Child Psychology and Psychiatry*, 58(7), 770–778. <https://doi.org/10.1111/jcpp.12694>.
- Harms, M. B., Bowen, K. E. S., Hanson, J. L., & Pollak, S. D. (2018). Instrumental learning and cognitive flexibility processes are impaired in children exposed to early life stress. *Developmental Science*, 21(4), e12596. <https://doi.org/10.1111/desc.12596>.
- Heron-Delaney, M., Anzures, G., Herbert, J. S., Quinn, P. C., Slater, A. M., Tanaka, J. W., Lee, K., & Pascalis, O. (2011). Perceptual training prevents the emergence of the other race effect during infancy. *PLoS One*, 6(5), e19858. <https://doi.org/10.1371/journal.pone.0019858>.
- Hoemann, K., Xu, F., & Barrett, L. F. (2019). Emotion words emotion concepts and emotional development in children: A constructionist hypothesis. *Developmental Psychology*, 55(9), 1830–1849. <https://doi.org/10.1037/dev0000686>.
- Krogh, L., Vlach, H., & Johnson, S. P. (2013). Statistical learning across development: Flexible yet constrained. *Frontiers in Psychology*, 3. <https://doi.org/10.3389/fpsyg.2012.00598>.
- Meeren, H. K., van Heijnsbergen, C. C., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences*, 102(45), 16518–16523. <https://doi.org/10.1073/pnas.0507650102>.
- Mermier, J., Quadrelli, E., Turati, C., & Bulf, H. (2022). Sequential learning of emotional faces is statistical at 12 months of age. *Infancy*, 27(3), 479–491. <https://doi.org/10.1111/inf.12463>.
- Mitchel, A. D., & Weiss, D. J. (2014). Visual speech segmentation: Using facial cues to locate word boundaries in continuous speech. *Language, Cognition and Neuroscience*, 29(7), 771–780. <https://doi.org/10.1080/01690965.2013.791703>.
- Moses, L. J., Baldwin, D. A., Rosicky, J. G., & Tidball, G. (2001). Evidence for referential understanding in the emotions domain at twelve and eighteen months. *Child Development*, 72(3), 718–735. <https://doi.org/10.1111/1467-8624.00311>.
- Mumme, D. L., Fernald, A., & Herrera, C. (1996). Infants' responses to facial and vocal emotional signals in a social referencing paradigm. *Child Development*, 67(6), 3219–3237. <https://doi.org/10.1111/j.1467-8624.1996.tb01910.x>.
- Niedenthal, P. M., Rychlowska, M., & Wood, A. (2017). Feelings and contexts: Socioecological influences on the nonverbal expression of emotion. *Current Opinion in Psychology*, 17, 170–175. <https://doi.org/10.1016/j.copsyc.2017.07.025>.
- Patzwald, C., Curley, C. A., Hauf, P., & Elsner, B. (2018). Differential effects of others' emotional cues on 18-month-olds' preferential reproduction of observed actions. *Infant Behavior and Development*, 51, 60–70. <https://doi.org/10.1016/j.infbeh.2018.04.002>.
- Plate, R. C., Shutts, K., Cochrane, A., Green, C. S., & Pollak, S. D. (2021). Testimony bias lingers across development under uncertainty. *Developmental Psychology*, 57(12), 2150–2164. <https://doi.org/10.1037/dev0001253>.
- Pollak, S. D., Messner, M., Kistler, D. J., & Cohn, J. F. (2009). Development of perceptual expertise in emotion recognition. *Cognition*, 110(2), 242–247. <https://doi.org/10.1016/j.cognition.2008.10.010>.
- Ruba, A. L., Pollak, S. D., & Saffran, J. R. (2022). Acquiring complex communicative systems: Statistical learning of language and emotion. *Topics in Cognitive Science*. <https://doi.org/10.1111/tops.12612>.
- Ruba, A. L., & Repacholi, B. M. (2020). Beyond language in infant emotion concept development. *Emotion Review*, 12(4), 255–258. <https://doi.org/10.1177/1754073920931574>.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926–1928. <https://doi.org/10.1126/science.274.5294.1926>.
- Santolin, C., & Saffran, J. R. (2018). Constraints on statistical learning across species. *Trends in Cognitive Sciences*, 22(1), 52–63. <https://doi.org/10.1016/j.tics.2017.10.003>.
- Schapiro, A. C., Kustner, L. V., & Turk-Browne, N. B. (2012). Shaping of object representations in the human medial temporal lobe based on temporal regularities. *Current Biology*, 22(17), 1622–1627. <https://doi.org/10.1016/j.cub.2012.06.056>.
- Schirmer, A., & Adolphs, R. (2017). Emotion perception from face, voice, and touch: Comparisons and convergence. *Trends in Cognitive Sciences*, 21(3), 216–228. <https://doi.org/10.1016/j.tics.2017.01.001>.
- Sherman, B. E., Graves, K. N., & Turk-Browne, N. B. (2020). The prevalence and importance of statistical learning in human cognition and behavior. *Current Opinion in Behavioral Sciences*, 32, 15–20. <https://doi.org/10.1016/j.cobeha.2020.01.015>.
- Siegelman, N., Bogaerts, L., Christiansen, M. H., & Frost, R. (2017). Towards a theory of individual differences in statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1711), 20160059. <https://doi.org/10.1098/rstb.2016.0059>.
- Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language*, 81, 105–120. <https://doi.org/10.1016/j.jml.2015.02.001>.
- Siegelman, N., Bogaerts, L., Kronenfeld, O., & Frost, R. (2018). Redefining “learning” in statistical learning: What does an online measure reveal about the assimilation of visual regularities? *Cognitive Science*, 42, 692–727. <https://doi.org/10.1111/cogs.12556>.
- Smith, L. B., Jayaraman, S., Clerkin, E., & Yu, C. (2018). The developing infant creates a curriculum for statistical learning. *Trends in Cognitive Sciences*, 22(4), 325–336. <https://doi.org/10.1016/j.tics.2018.02.004>.
- Sorce, J. F., Emde, R. N., Campos, J. J., & Klinnert, M. D. (1985). Maternal emotional signaling: Its effect on the visual cliff behavior of 1-year-olds. *Developmental Psychology*, 21(1), 195–200. <https://doi.org/10.1037/0012-1649.21.1.195>.
- Tanaka, J. W., Curran, T., & Sheinberg, D. L. (2005). The training and transfer of real-world perceptual expertise. *Psychological Science*, 16(2), 145–151. <https://doi.org/10.1111/j.0956-7976.2005.00795.x>.
- Tanaka, J. W., Heptonstall, B., & Hagen, S. (2013). Perceptual expertise and the plasticity of other-race face recognition. *Visual Cognition*, 21(9–10), 1183–1201. <https://doi.org/10.1080/13506285.2013.826315>.
- Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology*, 23(3), 457–482. [https://doi.org/10.1016/0010-0285\(91\)90016-H](https://doi.org/10.1016/0010-0285(91)90016-H).

- Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., Marcus, D. J., Westerlund, A., Casey, B. J., & Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research, 168*(3), 242–249. <https://doi.org/10.1016/j.psychres.2008.05.006>.
- Van de Cruys, S., de-Wit, L., Evers, K., Boets, B., & Wagemans, J. (2013). Weak priors versus overfitting of predictions in autism: Reply to Pellicano and Burr (*TICS*, 2012). *I-Perception, 4*(2), 95–97. <https://doi.org/10.1068/i0580ic>.
- Van de Cruys, S., Evers, K., Van der Hallen, R., Van Eylen, L., Boets, B., de-Wit, L., & Wagemans, J. (2014). Precise minds in uncertain worlds: Predictive coding in autism. *Psychological Review, 121*(4), 649–675. <https://doi.org/10.1037/a0037665>.
- Walle, E. A., Reschke, P. J., & Knothe, J. M. (2017). Social referencing: Defining and delineating a basic process of emotion. *Emotion Review, 9*(3), 245–252. <https://doi.org/10.1177/1754073916669594>.
- Wu, Y., Muentener, P., & Schulz, L. E. (2017). One- to four-year-olds connect diverse positive emotional vocalizations to their probable causes. *Proceedings of the National Academy of Sciences, 114*(45), 11896–11901. <https://doi.org/10.1073/pnas.1707715114>.
- Wu, Y., & Gweon, H. (2021). Preschool-aged children jointly consider others' emotional expressions and prior knowledge to decide when to explore. *Child Development, 92*(3), 862–870. <https://doi.org/10.1111/cdev.13585>.
- Wu, Y., Schulz, L. E., Frank, M. C., & Gweon, H. (2021). Emotion as information in early social learning. *Current Directions in Psychological Science*. Advance Online Publication., 30, 468–475. <https://doi.org/10.1177/09637214211040779>.
- Zebrowitz, L. A. (2017). First impressions from faces. *Current Directions in Psychological Science, 26*(3), 237–242. <https://doi.org/10.1177/0963721416683996>.