Received: 19 July 2021 | Accepted: 7 February 2022

DOI: 10.1111/infa.12463

RESEARCH ARTICLE

INFANCY THE OFFICIAL JOURNAL OF THE INTERNATIONAL CONGRESS OF INFANT STUDIES WILEY

Sequential learning of emotional faces is statistical at 12 months of age

Julia Mermier[1] | **Ermanno Quadrelli**[1,2] | **Chiara Turati**[1,2] | **Hermann Bulf**[1,2]

[1]Department of Psychology, University of Milano-Bicocca, Milano, Italy

[2]NeuroMI, Milan Center for Neuroscience, Milano, Italy

Correspondence
Julia Mermier, Department of Psychology, University of Milano-Bicocca, Piazza dell'Ateneo Nuovo 1, 20126, Milano, Italy.
Email: julia.mermier@unimib.it

Funding information
H2020 Marie Skłodowska-Curie Actions, Grant/Award Number: 765298

Abstract

Infants are capable of extracting statistical regularities from continuous streams of elements, which helps them structuring their surrounding environment. The current study examines 12-month-olds' capacity to extract statistical information from a sequence of emotional faces. Using a familiarization procedure, infants were presented with videos of two actresses expressing the same facial emotion, and subsequently turning toward or away from each other. Videos displayed different emotions (i.e., anger, happiness, fear, sadness, surprise, amusement, disgust, and exasperation) and were organized sequentially, so that the transitional probabilities between videos were highly predictable in some cases, and less predictable in others. At test, infants discriminated highly predictable from low predictable transitional probabilities, suggesting that they extracted statistical regularities from the sequence of emotional faces. However, when examining the *looking toward* and the *looking away* conditions separately, infants showed evidence of statistical learning in the *looking toward* condition only. Together, these findings suggest that 12-month-old infants rely on statistical learning to segment a continuous sequence of emotional faces, although this ability can be modulated by the nature of the stimuli. The contribution of statistical learning to structure infants' social environment is discussed.

1 | INTRODUCTION

Our understanding of others' behaviors and mental states highly depends on a successful decoding of emotions. In particular, facial expressions of emotions play a crucial role in social interactions, especially for pre-verbal infants, who rely mainly on non-linguistic cues to communicate. Observing their surrounding environment, infants constantly witness social interactions in which facial expressions of emotions succeed one another. These emotional expressions typically show some regularities in their order of presentation, with some emotions being regularly expressed consecutively, to form a coherent sequence. Indeed, the specific order in which emotions are displayed reflects the mental states of its protagonist and contributes to determine the overall meaning of the interaction. For example, if during a social interaction, individual 1 expresses anger toward individual 2, and individual 2 likewise responds with anger, the interaction is likely to turn to conflict. Instead, individual 2 responding with fear rather than anger signals that he or she feels threatened, and that the interaction might be interpreted differently, maybe suggesting that the two protagonists have unequal power within the relationship. Thus, being able to track regularities in sequences of facial expressions of emotions could potentially help infants structuring observed social interactions and facilitate their comprehension of the emotional context, as well as the nature of the social interactions. However, whether infants possess the ability to detect coherent patterns embedded in a sequence of different emotional faces has yet to be determined.

In other domains, previous research has shown that infants are able to extract predictable statistical regularities from a complex, continuous stream of elements. This capacity, referred to as statistical learning (SL), appears to rely on the transitional probabilities (TP) of a sequence of elements, that is, the probability of an item X to be followed by an item Y (Saffran et al., 1996). It is thought to appear early on in development (Bulf et al., 2011; Teinonen et al., 2009) and to support infants' comprehension and learning of relationships between the elements of a sequence, allowing the prediction of its upcoming stimuli (Kirkham et al., 2007).

Statistical learning was first evidenced in the linguistic domain, where 8-month-olds were able to segment words from a fluent artificial speech, using differences in TPs as the only cue for segmentation (Aslin et al., 1998; Saffran et al., 1996). Likewise, SL was found to be at the basis of infants' capacity to track TPs in non-linguistic auditory sequences (Hannon & Johnson, 2005; Saffran et al., 1999) and visual sequences (Bulf et al., 2011; Kirkham et al., 2002, 2007). Moreover, it was demonstrated that infants can apply SL to many different categories of stimuli, such as animated objects (Stahl et al., 2014), human gestures (Quadrelli et al., 2020; Roseberry et al., 2011), and complex human actions (Meyer et al., 2011; Monroy et al., 2017, 2019; Saylor et al., 2007). Taken together, these findings suggest that SL is a domain-general mechanism (for a review see Saffran & Kirkham, 2018) that might operate similarly on different categories of stimuli, supporting infants' comprehension of the surrounding environment.

Interestingly, infants' capacity to detect regularities from a sequence of elements seems to be modulated by the content of the stimuli, and, in particular, by the presence of social signals (i.e., faces). For example, in their study, Bulf et al. (2015) presented 7-month-olds with photographs of neutral faces of different identities, whose order of presentation followed a specific repetition-based, rule-like pattern (i.e., ABB or ABA). Infants were capable of learning the rule of presentation and generalize it to new face identities when the faces were presented in an upright, but not inverted, configuration. Moreover, the affective content of the faces was shown to affect this capacity, which was maintained when identities displayed happy expressions, but was disrupted when identities displayed angry expressions (Quadrelli et al., 2019). These studies demonstrate that infants' ability to detect repetition-based regularities from sequences of facial emotions and identities is affected by the social content of the stimuli that make up the sequence.

Previous research has also demonstrated that the social context in which emotions are expressed considerably influences infants' behavior and learning. For instance, Montague and Walker-Andrews (2001) demonstrated that, when displayed in the familiar context of a peekaboo game, 4-month-olds' discrimination of emotions (happy/surprise vs. anger/fear/sadness) was facilitated. Furthermore, 18-month-olds were more prone to copy the specific action of a model that looked engaging and social (i.e., smiling and eye-contact), than a model that seemed aloof and disinterested (Nielsen, 2006). Besides, the mere presence of salient social cues was shown to greatly influence infants' behavior and learning. For instance, 6-month-old infants preferred to attend to a typical face-to-face rather than an unconventional back-to-back social interaction, and made more gaze shifts between the two actors when they were facing each other (Augusti et al., 2010). Likewise, associative learning in 13-month-old infants was reinforced when target videos displayed a social interaction as compared to non-interactive control conditions (Thiele et al., 2021).

Taken together, the existing literature emphasizes the fundamental role of social cues on infants' perception and understanding of others' actions and interactions, and their significant influence on broader learning processes. Nonetheless, it remains unknown whether infants' ability to detect regularities from their social environment extends to situations in which the sequence of social stimuli (e.g., faces) is statistically defined, and the only available cue is the TP between stimuli. Specifically, no study has previously investigated whether infants are able to track statistical regularities from a predictable sequence of emotional faces. Furthermore, in a natural environment, emotional facial expressions are more frequently embedded within contexts of social interactions involving at least two protagonists, rather than single isolated faces. The social context in which emotions are expressed can potentially influence the meaning of the interaction, as well as infants' behavior and learning.

The primary aim of our study was to investigate whether infants are capable of extracting statistical regularities from a sequence of emotional faces with differences in TPs as the only cue for segmentation. As a secondary aim, we also examined whether the degree of sociality of the stimuli modulates this ability. To this end, we presented infants with a sequence of videos representing two actresses expressing eight different emotional expressions: anger, happiness, fear, sadness, surprise, amusement, disgust, and exasperation. Emotions were selected for their very distinct perceptual properties in order to be easily encoded and based on past research examining infants' emotion recognition abilities. Indeed, by 7 months of age, infants were shown to discriminate anger, happiness, fear, sadness, surprise, and disgust (Farroni et al., 2007; Kotsoni et al., 2001; Ludemann & Nelson, 1988; Ruba et al., 2017), as well as different levels of intensity within a same emotion (e.g., mild versus extreme happiness and mild vs. extreme fear; Kuchuk et al., 1986; Ludemann & Nelson, 1988). Thus, in addition to the six basic emotions, we chose to present expressions of amusement (i.e., mild happiness) and exasperation. In each of the eight videos, the two actresses first faced forward with a neutral expression, then expressed the same facial emotion, and subsequently turned either toward (*looking toward* condition) or away (*looking away* condition) from each other while maintaining the expression at its peak. The videos of the *looking toward* condition thus contained highly salient social cues, as the two actresses faced and looked at each other, while the *looking away* condition contained less salient social cues, as the two actresses looked away from each other (see Augusti et al., 2010 for a similar manipulation). As infants' SL abilities were shown to be limited by their attentional and memory resources (Bulf et al., 2011; Vlach & Johnson, 2013), we decided to test 12-month-olds, who are older than usual SL studies' participants (Aslin et al., 1998; Kirkham et al., 2002; Roseberry et al., 2011; Stahl et al., 2014), presuming that they possessed the cognitive resources necessary for the statistical learning of the complex sequence that was shown.

During the familiarization phase, infants were presented with a sequence organized in four fixed pairs of videos, each pair called unit, and arranged so that the TP between videos was higher within each pair (units) than across pairs (part units). As units and part units occurred with the same frequency during familiarization, TPs were the only cue for extracting the statistical structure of the sequence. Indeed, transitions between two videos within a unit occurred with a TP of 1.0, while transitions between two videos across units occurred with a TP of 0.5. During the test phase, units and part units were presented in alternation, and infants' ability to discriminate them was assessed by measuring their looking times. We expected that 12-month-old infants would succeed in extracting the statistical regularities from the familiarization sequence, and thus look longer to the part units than to the units, as found in previous infant studies using the same method (e.g., Stahl et al., 2014). In addition, we explored whether the degree of sociality of the stimuli might affect infants' capacity to extract statistical regularities.

2 | METHODS

2.1 | Participants

Thirty-six healthy, full-term 12-month-old infants (19 females; M = 12.76 months, SD = 19 days, min age = 11.84, max age = 13.94) were included in the final sample. Half of the participants (N = 18) were randomly assigned to the *looking toward* condition and the other half (N = 18) to the *looking away* condition. Sixteen additional participants were tested but excluded from the final sample, due to parental interference (N = 2), because they did not watch enough stimuli (N = 12), or because looking times in at least one test trial exceeding ± 2.5 standard deviations (SD) from the overall group mean (N = 2; Johnson & Aslin, 1996; Koechlin, 1997). Based on existing literature using a similar procedure (e.g., Quadrelli et al., 2020; Roseberry et al., 2011; Saylor et al., 2007) and on an a priori power analysis performed using G*Power (Faul et al., 2007), a sample size of 32 participants was estimated in order to have 80% probability to detect a significant interaction (α = 0.05) with a medium effect size (r = .25), following Cohen's guidelines (Cohen, 1992), in a repeated measures ANOVA. Participants were recruited from birth records of neighboring cities via written invitation. The present study was conducted according to guidelines laid down in the Declaration of Helsinki, with written informed consent obtained from a parent or guardian for each child before any assessment or data collection. All procedures involving human subjects in this study were approved by the ethics committee at the University of Milano-Bicocca (Protocol number: 421).

2.2 | Apparatus

The study took place in a testing booth isolated from external noise and light. Infants sat on their parent's lap, at a distance of about 60 cm from a 24-inch computer screen with a resolution of 1600 x 1200 pixels. The stimuli were displayed on the computer screen using E-Prime 2.0. Parents were instructed not to interact in any way with the infant throughout the entire experimental procedure. A video camera hidden over the computer screen recorded the infant's face during the whole experiment and fed into a digital video recorder and a TV monitor, both located on the other side of the testing booth, out of sight of the participant. A trained experimenter, blind to the stimulus sequence and assigned condition, observed the live video displayed on the TV monitor and performed the online coding of infants' looking times by pressing the computer mouse when the infant was looking at the screen. A second experimenter further carried out the offline coding of the looking times by examining the recording of the infant's face frame by frame. A Pearson correlation between the online and offline coding was computed on the total fixation times during test trials, resulting in an inter-observer agreement of $r = .99$, $p < .01$.

2.3 | Stimuli

Stimuli consisted of eight colored videos in which two Caucasian females first faced forward with a neutral expression, then simultaneously expressed the same facial emotion, and subsequently turned either toward (*looking toward* condition) or away (*looking away* condition) from each other while maintaining the expression at its peak (Figure 1). The videos displayed eight different emotions: anger, happiness, fear, sadness, surprise, disgust, amusement, and exasperation, and were chosen to be as naturalistic as possible (i.e., dynamic, involving two identities, and with uncropped faces). Stimuli had a duration of 2000 ms and were created using the software Adobe Premiere Pro CC 2019 and Kinovea.

2.4 | Procedure

Stimuli were embedded in a frame with a visual angle that subtended 11° x 20° at a viewing distance of 60 cm. The eight videos were presented in a sequence and organized into four fixed pairs of videos called units (Figure 2). The four units were the same for all participants, so that the first video of a unit was always followed by the same second video. Thus, within units, the transitional probability (TP) between the first and second video was of 1.0. Similar to previous studies (Stahl et al., 2014), during the familiarization phase, two of the units were presented 10 times (i.e., high-frequency units), and the other two were presented 5 times (i.e., low-frequency units), for a total of 30 units. The units were displayed in a random order, with the constraint that one unit could not be followed by itself. The last video of a high-frequency unit and the first video of a low-frequency unit formed the part units and had an internal TP of 0.5. Part units occurred the same amount of time as low-frequency units (Aslin et al., 1998).

An animated attention getter was presented at the beginning of the familiarization phase, to attract the infant's attention toward the screen. Once the infant looked at the screen, the experimenter started the stimulus presentation. During familiarization, anytime the infant looked away from the screen for more than 1 s, the stimulus presentation was interrupted, and an audiovisual attention getter was displayed. As soon as the infant looked back to the screen, stimulus

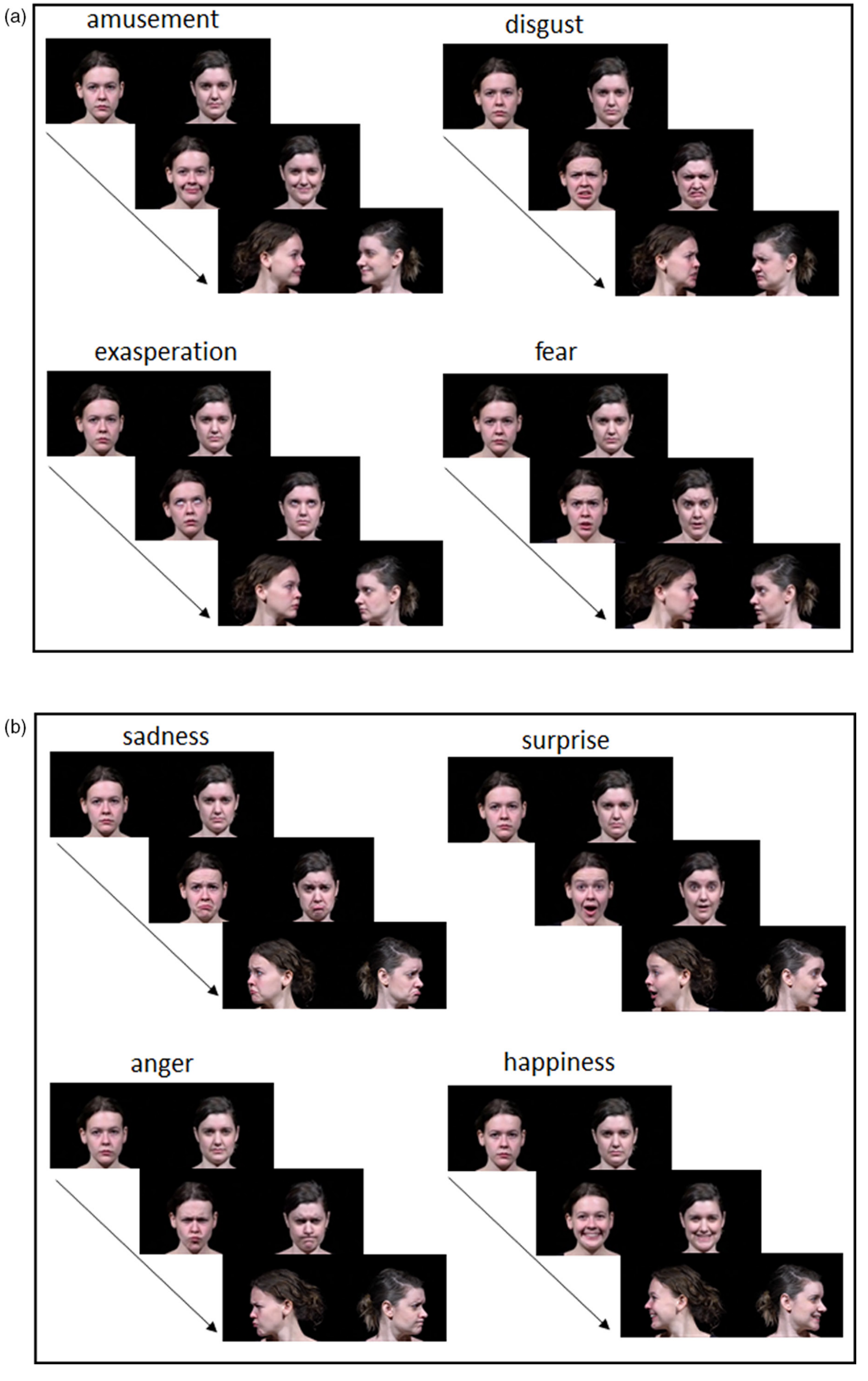

FIGURE 1 Schematic representation of the 8 emotional facial expressions presented in the task. The top part of the figure (a) represents four examples of emotions within the *looking toward* condition and the bottom part of the figure (b) represents four examples of emotions within the *looking away* condition

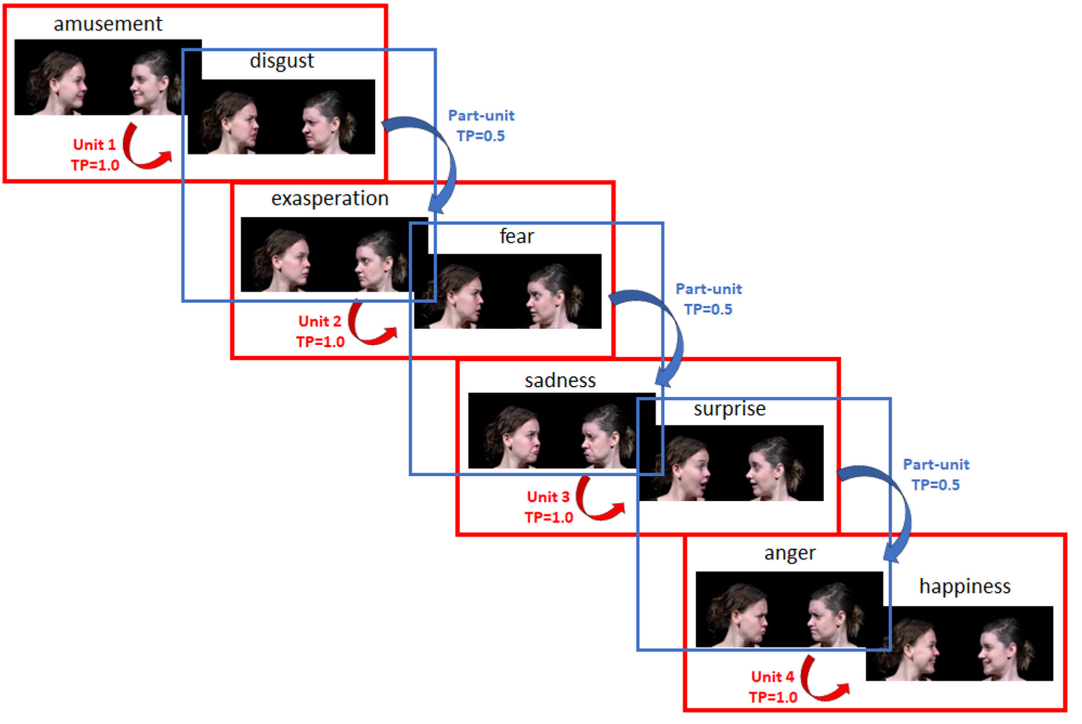

FIGURE 2 Example of stimuli presented in the task for the *looking toward* condition. In the *looking away* condition, the two actresses turn away from each other instead of turning towards each other. Videos are organized in 4 pairs of facial emotions called units, with a transitional probability of 1 within each unit (highly predictable transitions), and a transitional probability of 0.5 between units (low predictable transitions)

presentation resumed from where it had stopped. In this way, we ensured that all infants watched the entire familiarization. As each video lasted 2000 ms and was followed by a 200 ms interstimulus interval, the total duration of the familiarization phase was of 132 s. A 500 ms blank was displayed between the familiarization and the test phase.

The test phase consisted of eight test trials divided into two blocks, each block being composed of 2 low-frequency units and two part units trials in alternation. As low-frequency units and part units occurred with the same frequency during familiarization, the only cue allowing the discrimination between units and part units was their differing TPs. The test trials of a block were displayed in a loop, until the infants looked away for 1000 ms, or for a maximum duration of 22 s. Before each block, an attention getter was presented until the infant looked at the center of the screen. The order of the test trials (i.e., unit vs. part unit first) was counterbalanced among infants, and the same procedure was applied for both the *looking toward* and *looking away* conditions.

2.5 | Data analysis

A preliminary inspection of our data indicated that raw looking times at test were not normally distributed for both units and part units (Ws > 0.86, *ps* < 0.05). As a consequence, data were logarithmically transformed for the statistical analyses, following Csibra et al. (2016) recommendations.

We conducted a repeated measures analysis of variance (ANOVA) on logarithmically transformed looking times at test, with Block (first vs. second) and Trial type (units vs. part units) as within-subject factors and Condition (*looking toward* vs. *looking away*) and Trial order (units first vs. part units first) as between-subject factors. Planned comparisons were also conducted to explore infants' looking time patterns separately for both the *looking toward* and *looking away* conditions, by performing repeated measures ANOVAs on logarithmically transformed looking times at test, with Block (first vs. second) and Trial type (units vs. part units) as within-subject factors and Trial order (units first vs. part units first) as between-subject factors. These comparisons were planned a priori, based upon the existing literature and our predictions that SL would be facilitated in the *looking toward* condition, and impaired in the *looking away* condition. Pairwise comparisons were performed by applying t-tests and the Fisher's least significant difference procedure (Howell, 2012), and Holm–Bonferroni correction was used where appropriate (Abdi, 2010). The Greenhouse–Geisser correction for non-sphericity was used to adjust degrees of freedom as appropriate. Effect sizes were estimated using the η^2_p measure, and the data are reported as means and standard deviations (SDs). All statistical analyses were performed on Jamovi 1.6.15 (https://jamovi.org) using a two-tailed 0.05 level of significance. Following recent recommendations on best practices in infant looking-time research (Oakes, 2017), in order to strengthen our results, we also performed Bayesian analyses by using the default Cauchy prior (r = .707). Using the Jamovi formalism, the index next to the Bayes Factors (BF) indicates that the null hypothesis (H_0) is in the denominator and the alternative hypothesis (H_1) is in the numerator. Thus, BF_{10} is $p(\text{data}|H_1)/p(\text{data}|H_0)$, with $BF_{10} > 10$ considered as strong evidence for an effect, and $3 < BF_{10} < 10$ considered as moderate evidence.

3 | RESULTS

The main repeated measures ANOVA revealed a significant main effect of Block $F(1,32) = 7.40$, $p = .010$, $\eta^2_p = 0.188$, and a significant main effect of Trial Type, $F(1,32) = 7.28$, $p = .011$, $\eta^2_p = 0.185$. Infants looked significantly longer to the first block (M = 63.17 s, SD = 23.38) than to the second block (M = 55.07 s, SD = 21.29) and significantly longer to part units (M = 62.89 s, SD = 21.03) than units (M = 55.34 s, SD = 23.89; Figure 3). Examination of the data for individual infants through binomial tests confirmed the results of the analysis on looking times, revealing that 25 out of the 36 twelve-month-old infants looked longer to the part units compared with the units (25 vs. 11, p = .029). Two-tailed paired sample Bayesian t-tests confirmed the results obtained from frequentist analysis, showing moderate evidence for a difference between units and part units ($BF_{10} = 4.27$) and moderate evidence for a difference between the first and the second block ($BF_{10} = 5.22$). The difference in looking times between the two blocks likely reflects a decrease in infants' attention during the second part of the study, which is commonly found in looking time paradigms using long familiarization phases. No other main or interaction effects reached statistical significance (all $ps > .09$).

Despite the absence of main effect or interaction of the factor Condition, we had planned to examine the *looking toward* and *looking away* conditions separately. The ANOVA performed on the *looking toward* condition revealed a significant main effect of Block $F(1,16) = 5.39$, $p = .034$, $\eta^2_p = .252$, and a significant main effect of Trial Type, $F(1,16) = 6.92$, $p = .018$, $\eta^2_p = .302$. Infants looked significantly longer to the first block (M = 63.40 s, SD = 25.65) than to the second block (M = 54.35 s, SD = 23.31), and significantly longer to part units (M = 63.91 s, SD = 25.14) than

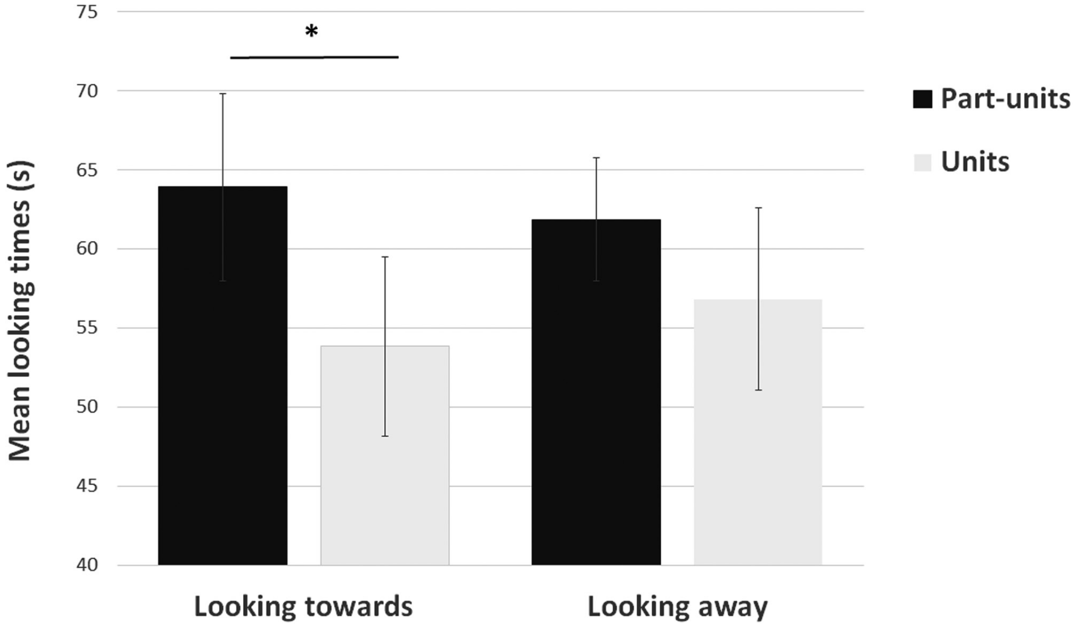

FIGURE 3 Mean looking times (±SE) to units and part units for the *looking toward* and the *looking away* conditions. Infants looked longer at the part units than at the units only in the *looking toward* condition. *$p < .05$

units ($M = 53.84$ s, SD = 23.99) (Figure 3). Two-tailed paired sample Bayesian *t*-tests confirmed the results obtained from frequentist analysis, showing moderate evidence for a difference between units and part units ($BF_{10} = 3.96$) and anecdotal evidence for a difference between the first and the second block ($BF_{10} = 2.08$). No other main or interaction effects attained statistical significance (all $ps > .32$). The ANOVA performed on the *looking away* condition did not reveal any main or interaction effects (all $ps > .12$). Two-tailed paired sample Bayesian *t*-tests further confirmed the lack of significant results resulted from the frequentist analysis, showing anecdotal evidence for the lack of difference between units ($M = 56.84$ s, SD = 24.40 s) and part units ($M = 61.87$ s, SD = 16.62 s) in the *looking away* condition ($BF_{10} = 0.57$; Figure 3).

4 | DISCUSSION

The current study sought to determine whether 12-month-old infants are capable of extracting statistical regularities from a sequence of interacting emotional faces, with differences in TPs as only cue for the segmentation. In addition, it investigated whether the degree of sociality of the stimuli modulates this ability. Results demonstrated that infants are indeed able to segment a continuous sequence of emotional faces relying solely on the differences in TPs between stimuli. These findings are in line with previous work showing that infants can apply SL to different categories of stimuli (e.g., auditory sequences, Hannon & Johnson, 2005; complex human actions, Monroy et al., 2017; human gestures, Quadrelli et al., 2020), extending it to sequences of emotional faces. They further support the view that SL is based on the same learning mechanisms across domains and modalities (Saffran & Kirkham, 2018) and allows infants to structure their surrounding environment, facilitating its comprehension. Future studies should investigate whether all emotions have the same impact on infants' ability to track statistics from a sequence of faces or whether some type of emotions or emotional pairs (e.g., the co-occurrence of emotional faces with positive vs. negative valence) might differently affect infants' learning abilities.

Importantly, the current study does not only extend the range of application of infants' SL, but also informs about its potential underlying mechanism. Indeed, as our design ensured that units and part units occurred with the same frequency during familiarization, the only cue upon which infants could base their segmenting of the sequence of interacting emotional faces was the differing TPs. Together with previous research examining SL of other categories of stimuli (e.g., words, Aslin et al., 1998; gestures, Quadrelli et al., 2020; events, Stahl et al., 2014), our findings suggest that the computing of TPs might be a common mechanism underlying SL across domains and modalities. As the functioning of infants' SL appears to be strongly affected by the type of input to be learned (Krogh et al., 2013), future studies should examine infants' SL abilities with emotional stimuli from different domains, as for example emotional vocal expressions or emotional gestures.

Although no evidence for a difference between conditions was found in the main analysis, the examination of the *looking toward* and *looking away* conditions separately nevertheless hinted at potential differences in looking time patterns. Indeed, we found evidence for a difference between units and part units in the *looking toward* condition, but not in the *looking away* condition, even if the trend in infants' looking time patterns was similar in the two conditions. This suggests that the SL of a sequence of emotional faces might be marginally influenced by the degree of sociality of the stimuli, with highly salient social stimuli such as those displayed in the *looking toward* condition promoting infants' SL, and less salient social stimuli such as those displayed in the *looking away* condition impairing it. This is in line with previous studies showing that social stimuli such as two actors facing or interacting with each other, as compared to actors sitting back-to-back, promoted infants' SL of a gestures sequence (Quadrelli et al., 2020) and helped binding actions into a collaborative sequence (Fawcett & Gredebäck, 2013). Alternatively, the lack of evidence for SL in the *looking away* condition might be based on the familiarity, rather than the sociality of the stimuli. Indeed, it is quite uncommon for infants to witness social interactions in which the two protagonists express an emotion, and then turn away from each other. Thus, the stimuli of the *looking away* condition might have been too unusual for infants and disrupted their ability to extract the statistical regularities from the sequence of emotions. In sum, the design and results of our study do not allow us to draw a firm conclusion about whether and why SL was disrupted in the *looking away* condition. Future research should further examine the role of the sociality of the stimuli in infants' SL, for example, by manipulating different social cues while keeping constant the familiarity of the stimuli. Besides, SL abilities only diverged when looking at both conditions separately, but did not significantly differ when comparing them directly. These findings thus require replication to be confirmed, and more research is needed to uncover the different factors influencing infants' SL of emotional faces. Indeed, in the current study, the stimuli of the *looking away* condition also contained several social cues. Even though the two actresses ended up looking away from each other, they started facing forward, looking toward the participants, and expressed the same emotion at the same time. Thus, infants could have interpreted the actresses' initial emotional expression (i.e., when facing forward) as directed to themselves, and their subsequent synchronicity of emotional behaviors as a social marker, considering the two actresses as acting jointly, rather than individually. Furthermore, infants' high familiarity (Gebhart et al., 2009) and perceptual expertise (Saffran et al., 2007) of faces, which are inherently social and highly salient per se, and the fact that our stimuli were presented in a very naturalistic way (i.e., dynamic and uncropped faces), similarly to what is observed in everyday life, might have promoted infants' learning. These facilitating factors, together with infants' increasing cognitive capacities at the end of the first year, might have decreased the possibility to find a more pronounced difference between the looking forward and looking away conditions. Future research could further clarify this issue investigating infants' SL using conditions that are more markedly contrasted in terms of level of sociality.

In conclusion, the current study presents evidence that 12-month-old infants are capable of extracting statistical regularities from a highly complex sequence of emotional faces, using TPs as only cue for segmentation. Thus, SL seems to be a crucial mechanism, which allows infants to structure their surrounding environment through the detection of statistical regularities (Saffran, 2018), providing a foundation for its comprehension. One essential aspect of infants' comprehension of their environment is the understanding of the social interactions surrounding them. These social interactions often consist in a succession of facial expressions of emotions, following specific patterns of occurrence, which determine the meaning of the interaction and reflect the mental states of its protagonists. The present study demonstrates that 12-month-old infants are capable of extracting these specific patterns of occurrence, in order to structure the social interactions they observe. This structuring might be the first step toward a deeper comprehension of social interactions; nevertheless, further research is needed to uncover whether it could be a potential foundation for more complex reasoning such as the understanding of the protagonists' mental states (e.g., Saylor et al., 2007).

ACKNOWLEDGMENTS

The authors are grateful to all families who kindly agreed to participate in this study. This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 765298. The authors declare no conflicts of interest with regard to the funding source for this study. Open Access funding provided by Universita degli Studi di Milano-Bicocca within the CRUI-CARE Agreement. [Correction added on 2 Jun 2022, after first online publication: CRUI funding statement has been added.]

AUTHOR STATEMENT

The work presented in this manuscript has not been published and is not under consideration for publication elsewhere. We certify that all authors of the manuscript have agreed to the listing and have seen and approved the revised version of the manuscript. We also certify that there are no affiliations with or involvement in any organization or entity with a direct financial interest in the subject matter or materials discussed in the manuscript. Research was conducted in accordance with APA ethical standards in the treatment of the study sample. Parents gave written informed consent for their infants' participation.

ORCID

Julia Mermier https://orcid.org/0000-0002-7680-9469
Ermanno Quadrelli https://orcid.org/0000-0002-6834-0912
Chiara Turati https://orcid.org/0000-0001-7496-1830
Hermann Bulf https://orcid.org/0000-0003-4121-1341

REFERENCES

Abdi, H. (2010). Holm's sequential Bonferroni procedure. *Encyclopedia of Research Design*, *1*(8), 1–8.

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, *9*(4), 321–324. https://doi.org/10.1111/1467-9280.00063

Augusti, E.-M., Melinder, A., & Gredebäck, G. (2010). Look who's talking: Pre-verbal infants' perception of face-to-face and back-to-back social interactions. *Frontiers in Psychology*, *1*, 161. https://doi.org/10.3389/fpsyg.2010.00161

Bulf, H., Brenna, V., Valenza, E., Johnson, S. P., & Turati, C. (2015). Many faces, one rule: The role of perceptual expertise in infants' sequential rule learning. *Frontiers in Psychology*, *6*, 1595. https://doi.org/10.3389/fpsyg.2015.01595

Bulf, H., Johnson, S. P., & Valenza, E. (2011). Visual statistical learning in the newborn infant. *Cognition*, *121*(1), 127–132. https://doi.org/10.1016/j.cognition.2011.06.010

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science*, *1*(3), 98–101. https://doi.org/10.1111/1467-8721.ep10768783

Csibra, G., Hernik, M., Mascaro, O., Tatone, D., & Lengyel, M. (2016). Statistical treatment of looking-time data. *Developmental Psychology*, *52*(4), 521–536. https://doi.org/10.1037/dev0000083

Farroni, T., Menon, E., Rigato, S., & Johnson, M. H. (2007). The perception of facial expressions in newborns. *European Journal of Developmental Psychology*, *4*(1), 2–13. https://doi.org/10.1080/17405620601046832

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fawcett, C., & Gredebäck, G. (2013). Infants use social context to bind actions into a collaborative sequence. *Developmental Science*, *16*(6):841–849. https://doi.org/10.1111/desc.12074

Gebhart, A. L., Newport, E. L., & Aslin, R. N. (2009). Statistical learning of adjacent and nonadjacent dependencies among nonlinguistic sounds. *Psychonomic Bulletin & Review*, *16*(3), 486–490. https://doi.org/10.3758/PBR.16.3.486

Hannon, E. E., & Johnson, S. P. (2005). Infants use meter to categorize rhythms and melodies: Implications for musical structure learning. *Cognitive Psychology*, *50*(4), 354–377. https://doi.org/10.1016/j.cogpsych.2004.09.003

Howell, D. C. (2012). *Statistical methods for psychology* (8th ed.). Wadsworth Cengage Learning.

Johnson, S. P., & Aslin, R. N. (1996). Perception of object unity in young infants: The roles of motion, depth, and orientation. *Cognitive Development*, *11*(2), 161–180. https://doi.org/10.1016/S0885-2014(96)90001-5

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, *83*(2), B35–B42. https://doi.org/10.1016/S0010-0277(02)00004-5

Kirkham, N. Z., Slemmer, J. A., Richardson, D. C., & Johnson, S. P. (2007). Location, location, location: Development of spatiotemporal sequence learning in infancy. *Child Development*, *78*(5), 1559–1571. https://doi.org/10.1111/j.1467-8624.2007.01083.x

Koechlin, E. (1997). Numerical transformations in five-month-old human infants. *Mathematical Cognition*, *3*(2), 89–104. https://doi.org/10.1080/135467997387425

Kotsoni, E., de Haan, M., & Johnson, M. H. (2001). Categorical perception of facial expressions by 7-month-old infants. *Perception*, *30*(9), 1115–1125. https://doi.org/10.1068/p3155

Krogh, L., Vlach, H. A., & Johnson, S. P. (2013). Statistical learning across development: Flexible yet constrained. *Frontiers in Psychology*, *3*, 598. http://doi.org/10.3389/fpsyg.2012.00598

Kuchuk, A., Vibbert, M., & Bornstein, M. H. (1986). The Perception of smiling and its experiential correlates in three-month-old infants. *Child Development*, *57*(4), 1054–1061. https://doi.org/10.2307/1130379

Ludemann, P. M., & Nelson, C. A. (1988). Categorical representation of facial expressions by 7-month-old infants. *Developmental Psychology*, *24*(4), 492–501. https://doi.org/10.1037/0012-1649.24.4.492

Meyer, M., Baldwin, D. A., & Sage, K. (2011). Assessing young children's hierarchical action segmentation. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *33*(33), 3156–3161.

Monroy, C., Meyer, M., Gerson, S., & Hunnius, S. (2017). Statistical learning in social action contexts. *PLoS One*, *12*(5), e0177261. https://doi.org/10.1371/journal.pone.0177261

Monroy, C., Meyer, M., Schröer, L., Gerson, S. A., & Hunnius, S. (2019). The infant motor system predicts actions based on visual statistical learning. *NeuroImage*, *185*, 947–954. https://doi.org/10.1016/j.neuroimage.2017.12.016

Montague, D. P. F., & Walker-Andrews, A. S. (2001). Peekaboo: A new look at infants' perception of emotion expressions. *Developmental Psychology*, *37*(6), 826–838. https://doi.org/10.1037/0012-1649.37.6.826

Nielsen, M. (2006). Copying actions and copying outcomes: Social learning through the second year. *Developmental Psychology*, *42*(3), 555–565. https://doi.org/10.1037/0012-1649.42.3.555

Oakes, L. M. (2017). Sample size, statistical power, and false conclusions in infant looking-time research. *Infancy*, *22*(4), 436–469. https://doi.org/10.1111/infa.12186

Quadrelli, E., Conte, S., Macchi Cassia, V., & Turati, C. (2019). Emotion in motion: Facial dynamics affect infants' neural processing of emotions. *Developmental Psychobiology*, *61*(6), 843–858. https://doi.org/10.1002/dev.21860

Quadrelli, E., Monacò, S., Turati, C., & Bulf, H. (2020). Social context influences infants' ability to extract statistical information from a sequence of gestures. *Infant Behavior and Development*, *61*, 101506. https://doi.org/10.1016/j.infbeh.2020.101506

Roseberry, S., Richie, R., Hirsh-Pasek, K., Golinkoff, R. M., & Shipley, T. F. (2011). Babies catch a break: 7- to 9-month-olds track statistical probabilities in continuous dynamic events. *Psychological Science*, *22*(11), 1422–1424. https://doi.org/10.1177/0956797611422074

Ruba, A. L., Johnson, K. M., Harris, L. T., & Wilbourn, M. P. (2017). Developmental changes in infants' categorization of anger and disgust facial expressions. *Developmental Psychology*, *53*(10), 1826–1832. https://doi.org/10.1037/dev0000381

Saffran, J. R. (2018). Statistical learning as a window into developmental disabilities. *Journal of Neurodevelopmental Disorders*, *10*(1), 35. https://doi.org/10.1186/s11689-018-9252-y

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

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*(1), 27–52. https://doi.org/10.1016/S0010-0277(98)00075-4

Saffran, J. R., & Kirkham, N. Z. (2018). Infant statistical learning. *Annual Review of Psychology*, *69*(1), 181–203. https://doi.org/10.1146/annurev-psych-122216-011805

Saffran, J. R., Pollak, S. D., Seibel, R. L., & Shkolnik, A. (2007). Dog is a dog is a dog: Infant rule learning is not specific to language. *Cognition*, *105*(3), 669–680. https://doi.org/10.1016/j.cognition.2006.11.004

Saylor, M. M., Baldwin, D. A., Baird, J. A., & LaBounty, J. (2007). Infants' on-line segmentation of dynamic human action. *Journal of Cognition and Development*, *8*(1), 113–128. https://doi.org/10.1080/15248370709336996

Stahl, A. E., Romberg, A. R., Roseberry, S., Golinkoff, R. M., & Hirsh-Pasek, K. (2014). Infants segment continuous events using transitional probabilities. *Child Development*, *85*(6), 1821-1826. https://doi.org/10.1111/cdev.12247

Teinonen, T., Fellman, V., Näätänen, R., Alku, P., & Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neuroscience*, *10*(1), 21. https://doi.org/10.1186/1471-2202-10-21

Thiele, M., Hepach, R., Michel, C., Gredebäck, G., & Haun, D. B. M. (2021). Social interaction targets enhance 13-month-old infants' associative learning. *Infancy*, *26*(3), 409–422. https://doi.org/10.1111/infa.12393

Vlach, H. A., & Johnson, S. P. (2013). Memory constraints on infants' cross-situational statistical learning. *Cognition*, *127*(3), 375–382. https://doi.org/10.1016/j.cognition.2013.02.015

How to cite this article: Mermier, J., Quadrelli, E., Turati, C., & Bulf, H. (2022). Sequential learning of emotional faces is statistical at 12 months of age. *Infancy*, 27, 479–491. https://doi.org/10.1111/infa.12463