

A general serial dependence among various facial traits: Evidence from Markov Chain and derivative of Gaussian

Jun-Ming Yu

Department of Psychology, Soochow University,
Suzhou, China



Haojiang Ying

Department of Psychology, Soochow University,
Suzhou, China



The human vision system can extract a stable representation of the always-changing visual world. However, the mechanism underlying such perceptual continuity remains unclear. A possible candidate is the serial dependence: visual perception of an object is positively biased toward the visual input from the recent past. Does the visual system use one pattern of serial dependence for general purposes? Or different patterns of serial dependence for different visual tasks? Because different social facial traits (e.g., trustworthiness and dominance) are dissociable, it is reasonable to assume that the perception of different facial characteristics would require different patterns of serial dependences. In this study, we examine the existence and the similarities of the serial dependence(s) in the evaluation of seven facial characteristics (i.e., attractiveness, trustworthiness, confidence, dominance, intelligence, age, and aggressiveness). The convergent evidence from conventional Derivative of Gaussian fitting and Markov Chain modeling demonstrated that (1) serial dependence exists in judgments of all seven social facial characteristic, (2) the serial dependences of them are highly similar, and (3) the serial dependence follows the efficient coding. Thus it is highly possible that there exists a general serial dependence mechanism for (at least high-level) vision processing. Moreover, we used the Markov Chain modeling to better describe the transitional pattern of serial dependence, which is a kind of Markov process. These findings may shed light on future works regarding serial dependence, as well as face perception.

Introduction

We live in an ever-changing world where the visual properties constantly fluctuate because of factors such as occlusion, signal noise, and eye movements. However, the visual system can form a stable and continuous representation of the always-changing visual environment (Lieberman, Fischer, & Whitney,

2014). One possible reason why our visual system can achieve such temporal stability is serial dependence: visual perception of the current stimuli is positively biased by the recent visual history. Previous studies have supported serial dependence with various stimuli and tasks, from the perception of basic visual features such as orientation (Fischer & Whitney, 2014), motion (Alais, Leung, & Van der Burg, 2017) to even high-level face perception. Specifically, serial dependence has been observed in perception of face identity, gender, age, and facial attractiveness (Lieberman et al., 2014; Taubert, Alais & Burr, 2016; Van der Burg, Rhodes & Alais, 2019; Xia, Leib, & Whitney, 2016; Clifford, Watson, & White, 2018). Based on the converging evidence, it is highly likely that there exists a general pattern of serial dependence among various visual tasks. If so, then the serial dependence should be fundamental and ubiquitous like adaptation (Webster, 2015). Consequently, the serial dependence should be highly similar across different visual tasks in terms of its pattern and magnitude. Moreover, the functional role of it would also be consistent among various visual tasks. However, these have not yet been tested.

It is also unclear whether serial dependence exists in the perception of social characteristics. Individuals can quickly and effortlessly make social inferences about personal characteristics and traits from faces. These social judgments play a significant role in our daily life: for example, the perceived dominance and trustworthiness of faces can influence leadership decisions (Ballew & Todorov, 2007). As with other perceptual processes, face perception of social characteristics remains relatively stable despite constant changes in visual inputs because of factors including eye movement, shadows, and occlusion; thus researchers called it the invariant aspect of face processing (Balas & Verdugo, 2018; Bruce & Young, 1986; Calder & Young, 2005). However, the mechanism underlying the temporal stability of facial trait perception remains poorly understood. Does serial dependence operate on other facial evaluation as it does on facial

Citation: Yu, J.-M., & Ying, H. (2021). A general serial dependence among various facial traits: Evidence from Markov Chain and derivative of Gaussian. *Journal of Vision*, 21(13):4, 1–13, <https://doi.org/10.1167/jov.21.13.4>.



attractiveness? If so, does the visual system use one pattern of serial dependence on facial evaluations for the general purpose, or different levels and patterns of serial dependence for different tasks?

According to previous studies on serial dependence, perception of the visual world is a statistically dependent process in a time series (Burr & Cicchini, 2014; Fischer & Whitney, 2014; Liberman et al., 2014). Specifically, the visual perception of the current object is attracted toward the perception of other stimuli from the recent past (Liberman et al., 2014). Previous studies tended to model the serial dependence by fitting the data with the derivative of Gaussian function (DoG) and describe the magnitude of the serial dependence with amplitude of DoG (α ; Manassi, Liberman, Chaney, & Whitney, 2017; Fritsche, Spaak, & Lange, 2020). However, such a method of modeling serial dependence can only tell the existence and the strength of the *biasing* but not how the previous perception *biases* the current perception. Hence, to further elucidate the mechanism and the functional role of serial dependence, recent studies have used other modeling methods (Bliss, Sun, & D’Esposito, 2017; Kok, Taubert, der Burg, Rhodes, & Alais, 2017). For instance, Cicchini and colleagues (2018) used the Kalman-filter in serial dependence modeling. The Kalman-filter model assumes that the response to the current stimulus is given by a weighted sum of the current and previous stimuli, which depicted the serial dependence (Cicchini et al., 2018). These methods effectively predict how the amplitude of serial dependence changes as the function of interstimulus changes and clearly show that serial dependence leads to decreases in response errors and faster responses.

Using different research paradigms and analyzing methods, studies testing various stimuli showed convergent results that the representations of the current trials were positively affected by that of the previous ones (Fischer & Whitney, 2014; Fritsche et al., 2020; Liberman et al., 2014). Therefore the serial dependence closely resembles the Markov process, where the current state is only dependent on the previous state (Hägström, 2002). Therefore it is reasonable to use the Markov Chain modeling, an essential part of the Markov process, to analyze the detailed biasing of the serial dependence via the transitional probability matrix (one of the many outputs of Markov Chain modeling). To be specific, as the transitional probability matrix shows the evolutionary trajectory of the Markov Chain (a discrete state-space Markov process), it can reveal the transitional distribution among ratings/states of judgments and thus can potentially show how a specific rating/state of judgments is transitioned from another rating/state. Thus the transitional probability matrix would tell more detailed information than DoG analysis. For instance, although the results of a DoG analysis would indicate that perceiving facial attractiveness of the current face would be positively

affected by a preceding highly attractive face, it cannot (at least directly) indicate how an extremely attractive face (e.g., ranked as 7 on a 1-7 scale) differentiates from a highly attractive face (e.g., ranked 6 on the same scale) on their impacts of the perception of the current trial. Therefore the Markov Chain modeling, together with other methods, would unveil the mechanism of serial dependence in a clearer fashion.

The serial dependence has been observed for various visual tasks that engaged dissociable neural mechanisms at different visual stages. However, the differences in experimental designs and the stimuli prevented us from a direct comparison among different kinds of visual tasks. A good candidate to do so might be face perception. Previous studies found that perception of facial characteristics is dissociable in terms of computational and neural mechanisms. For example, a two-dimensional model has been proposed (Oosterhof & Todorov, 2008) that demonstrates that perception of facial characteristics can be modeled on two dimensions: trustworthiness/valence and dominance. Recent studies also pointed out that attractiveness should also be added to the model of social inferences from faces and developed and validated a three-dimensional model involving approachability, dominance, and youthful-attractiveness (Sutherland, Oldmeadow, Santos, Towler, Michael Burt, & Young, 2013; Etcoff, Stock, Haley, Vickery, & House, 2011; Jones, DeBruine, Flake, Liuzza, Antfolk, Arinze, Ndukaihe, Bloxson Lewis, Foroni, & Willis, 2021). Studies in neuroscience also suggested that different dimensions of character traits may be processed via shared and dissociable neural mechanisms and circuits (Oosterhof & Todorov, 2008; Taubert et al., 2016; Balas & Verdugo, 2018). For instance, perception of attractiveness involves orbitofrontal cortex (O’Doherty, Winston, Critchley, Perrett, Burt, & Dolan, 2003), perception of trustworthiness involves amygdala (Todorov, Gobbini, Evans, & Haxby, 2007); however, perception of both traits involved the core face processing areas at the structural encoding stage (e.g., fusiform face area; Eimer, 2000; Haxby, Hoffman, & Gobbini, 2000; Kanwisher, McDermott, & Chun, 1997). Therefore, if the serial dependence(s) of the perception of facial traits are conducted at the core system of face processing (i.e., Occipital Face Area [OFA], Fusiform Face Area [FFA], and Superior Temporal Sulcus (STS); Haxby et al., 2000) or even at the early stage of visual processing (e.g., V4), then the serial dependence(s) of them shall be highly similar. On the other hand, if the serial dependences of the perception of facial traits are processed at the extended system of face processing, then there might be two or three patterns of serial dependences based on the dissociable computational and neural mechanisms.

Here, we aim to investigate the existence and similarity of serial dependences in judgments of seven

facial traits (attractiveness, trustworthiness, confidence, dominance, intelligence, age, and aggressiveness). We selected these seven facial traits because some previous studies used these facial traits to investigate the structure of facial evaluation (Jones et al., 2021; Oosterhof & Todorov, 2008; Sutherland et al., 2013; Ying & Chen, 2021). These seven traits have been the common traits studied in these studies using different modeling methods from 14 traits subtracted from 1134 written personal descriptions (Oosterhof & Todorov, 2008). Therefore these seven facial traits are good candidates for investigating face perception in social inference. Two methods were used here to measure the serial dependencies among those facial traits: the conventional DoG fitting, which has been frequently used in serial dependence measurement (Liberman et al., 2014; Manassi et al., 2017; Fritsche et al., 2020), and the Markov Chain modeling as a new analysis method. We then analyzed the similarity of serial dependencies among the transitional probability matrices of the seven facial traits using matrix correlation analysis. If the serial dependencies (measured by the amplitude α of DoG fitting, and the transitional probability matrix of Markov Chain modeling) of the facial traits were similar between each other, then it is reasonable to assume that a general pattern of serial dependence was used to achieve temporal stability for a general purpose. If not, then it is reasonable to hypothesize that different serial dependencies were used by different facial traits. Therefore we would further analyze the correlation pattern to test the structure of face evaluation in terms of serial dependence.

Methods

Participants

Thirty-two students (13 males and 19 females; mean age of 19.5 years) from Soochow University, with normal or corrected-to-normal vision, participated in this study. All of them were naïve to the purpose of the experiment and provided informed consent, with ethics approved by the Ethics Committee at Soochow University, China. At the beginning, we chose 30 as the sample size (considering the dramatic different in experimental design and the purpose of the current study, we adapted the sample size of previous studies; Fritsche et al., 2020; Ceylan, Herzog, & Pascucci, 2021). However, two more subjects signed in before we closed the registration, so we still tested them.

Stimuli

We used the 45 Chinese faces previously used in one study (Burns, Yang, & Ying, 2021). The 45 faces were

from three face databases with ethnic Chinese faces: the Nanyang Facial Emotional Expression Database (Yap, Chan, & Christopoulos, 2016), the Taiwanese Facial Expression Image Database (Chen & Yen, 2007), and an unnamed database used by Wang, Yao, and Zhou (2015). Only the internal region of each face is visible (cropped by an oval mask). Faces were presented in black and white with luminance equalized by the SHINE toolbox (Willenbockel, Sadr, Fiset, Horne, Gosselin, & Tanaka, 2010).

Apparatus

Visual stimuli were presented on a 22-inch ASUS PG278Q LCD monitor (spatial resolution at 2560×1440 pixels, refresh rate at 120 Hz; see Zhang, Li, Miao, He, Zhang, & Zhang, 2018 for details). The monitor was controlled by a computer (Linux OS) running Matlab R2016a (MathWorks) via Psychtoolbox (Brainard, 1997; Pelli, 1997). Participants sat in an adjustable chair with their chins resting on a chin rest that was placed at 53 cm away from the monitor, with each pixel subtended 0.025° on the screen.

Procedures

The general procedure for this experiment was adapted from previous experiments testing the individual ratings of facial attractiveness (Rhodes & Jeffery, 2006; Ying, Burns, Choo, & Xu, 2020). The experiment was composed of seven blocks, and participants were asked to rate the seven facial characteristics, in separate blocks with randomized orders. In each block, participants rated the 45 faces four times in randomized orders in terms of one specific facial characteristic, resulting in 180 trials within each block. One facial characteristic was only tested in one block. In each trial, after a 0.5- to 1-second interval, the test face appeared for 1 second (Figure 1). The faces were presented one at a time in the center of the screen for 1 second, and after its disappearance, participants were asked to rate these faces in terms of the facial characteristic specified in the block on a 1-7 Likert scale (e.g., 1 for least attractive and 7 for most attractive). We

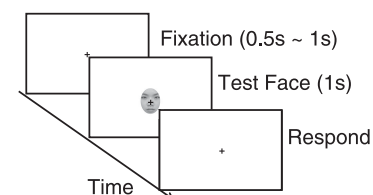


Figure 1. The trial sequence of the experiment.

used the Likert scale rating rather than adjust-to-match methods as in previous studies on purpose. Doing so allows us to model the responses by Markov Chain more easily (see the later section for more details). All test faces were displayed at the size of $3.28^\circ \times 4.23^\circ$.

Analysis

Two methods were used in measuring serial dependence in facial trait evaluation: the conventional modeling fitting and the Markov Chain modeling. The analysis was conducted with Matlab R2018a (MathWorks).

Conventional modeling fitting

Serial dependence of separate facial traits was first analyzed using conventional modeling fitting (Lieberman et al., 2014; Manassi et al., 2017; Fritsche et al., 2020). Judgment errors were computed as the difference between the participants' ratings and the estimated scores (calculated by averaging the judgments of all participants) of the target faces in terms of each facial trait. Judgment errors were then compared to the differences in scores of the facial traits between the current and previous trials (Lieberman et al., 2014). We pooled the judgment errors of all participants (excluding the first two trials, leaving 178 trials for fitting) and fitted the first DoG to the group data. The DoG was calculated by the function $y = xawce^{-(wx)^2}$, where parameter x is the difference in trait value between the current and 1-back target faces (1-back target face – current target face), a is half the peak-to-trough amplitude of the DoG, w scales the width of the DoG, and c is a constant ($\sqrt{2}/e^{-0.5}$) that scales the curve to make the a parameter equal to the peak amplitude. The amplitude parameter a was taken as the strength of the serial dependence bias, indicating the degree to which participants' judgments of each facial trait were biased towards the direction of the previous faces (Lieberman et al., 2014; Manassi et al., 2017; Fritsche et al., 2020). Here, we averaged the ratings of all participants for each facial trait and fitted the DoG function, which is a default protocol used by previous studies (Bliss et al., 2017; Fritsche et al., 2020; Ceylan et al., 2021) because it can systematically test serial dependence effect by measuring how the group average of response errors changes as a function of the difference between the previous and current facial trait value (Bliss et al., 2017; Fritsche et al., 2020; Ceylan et al., 2021). The value of parameter a of the DoG can reveal the level of serial dependence effect. To be specific, if participants' judgments of

the facial traits were systematically repelled or not influenced by the previous faces, then the parameter a of the DoG should be a negative value or even at zero, respectively. The width parameter w of the DoG curve was treated as a free parameter, constrained to a range of plausible values (w was set between 0 and 5, corresponding to the difference in facial trait value distributed between 0 to 5). We then fitted the Gaussian derivative using constrained nonlinear minimization of the residual sum of squares (Fischer & Whitney, 2014).

The permutation test was used to statistically evaluate the serial dependence biased on the group level, separately for each facial trait (based on Fritsche et al., 2020). A single permutation was conducted by shuffling the labels between the observed data (pooled from all participants), which generated an artificial null distribution of the “serial dependence” for each social facial characteristic. This permutation was repeated for 1000 times for each trait separately. The p value for each facial trait was calculated based on the z -score of the observed parameter a (also calculated based on the pooled data) on the null distribution of the given social facial characteristics separately. Note that, in this study, we analyzed the data with 1-back DoG which is the same as previous studies (Lieberman et al., 2014). We did try to fit the data with 2-back to 5-back DoG as well, but none of them reached significant value (in the permutation analysis). Thus we discarded these results. Moreover, this finding backed up our usage of first-order Markov Chain modeling in the later section.

Markov Chain modeling

In the present study, we used the Markov Chain modeling as a new analysis method to analyze serial dependence in facial trait judgments. Based on the previous studies on serial dependence, it is obvious that the perception of the current trial is heavily influenced by the previous trial (Lieberman et al., 2014; Xia et al., 2016), which fits the scope of the Markov Chain model. Therefore we used the Markov Chain model to summarize the participants' judgment of a particular facial trait and to further examine the serial dependence in facial trait judgment between current and previous trials. Briefly speaking, the Markov Chain is a time-series model that expresses transition between finite states according to certain transition probabilities. A Markov process with a state space of $\{1, 2, \dots, k\}$ can be determined by the k -order transition probability matrix (A), which specifies the transition probabilities between any two states. This right-stochastic transition matrix is able to describe the Markov Chain (i.e., a discrete state-space Markov process). Specifically, the participants' rating data of the

facial trait along with the progress of the experiment was regarded as time-series data, and the state space of the Markov model was defined as the rating range $\{1, 2, \dots, 7\}$. The transition information between the rating states was captured by the transition matrix, which contained the probabilities that participants' judgment of the facial trait transited among the seven rating states.

We pooled the data of all the participants and calculated the resulting transition matrix of each facial trait. The serial dependence of facial trait judgments could be measured by analyzing the distribution of the transitional probability matrix. More specifically, if ratings in the current trials were biased toward those of the previous trials, which indicated a classic serial dependence, then the transition matrix would be highly similar to a diagonal matrix (which indicates that the rating of a trial is strictly dependent on a previous trial). If ratings for the facial trait in the current trial were not influenced by a previous trial, then the transition matrix would be very unlikely to correlate with a diagonal matrix.

Results

Summary of raw ratings

First, we summarized the raw ratings from the participants (Figure 2). To test the extent to which the participants agreed on the ratings of each trait, Cronbach's alpha was computed. The ratings for the seven facial traits demonstrated high inter-rater reliability (Cronbach's $\alpha_{\text{attractiveness}} = 0.94$; Cronbach's $\alpha_{\text{trustworthiness}} = 0.91$; Cronbach's $\alpha_{\text{confidence}} = 0.90$; Cronbach's $\alpha_{\text{dominance}} = 0.88$; Cronbach's $\alpha_{\text{intelligence}} = 0.92$; Cronbach's $\alpha_{\text{age}} = 0.89$; Cronbach's $\alpha_{\text{aggressiveness}} = 0.91$). For each social characteristic, participants (in general) used the full range of the rating scale (*minimums* = 1, *maximums* = 7). With a repeated measures analysis of variance, we found that there are significant differences among the standard deviations (as an estimation of the degree of variability) of each participant's ratings (Greenhouse-Geisser corrected $F[4.23, 131.11] = 4.03$,

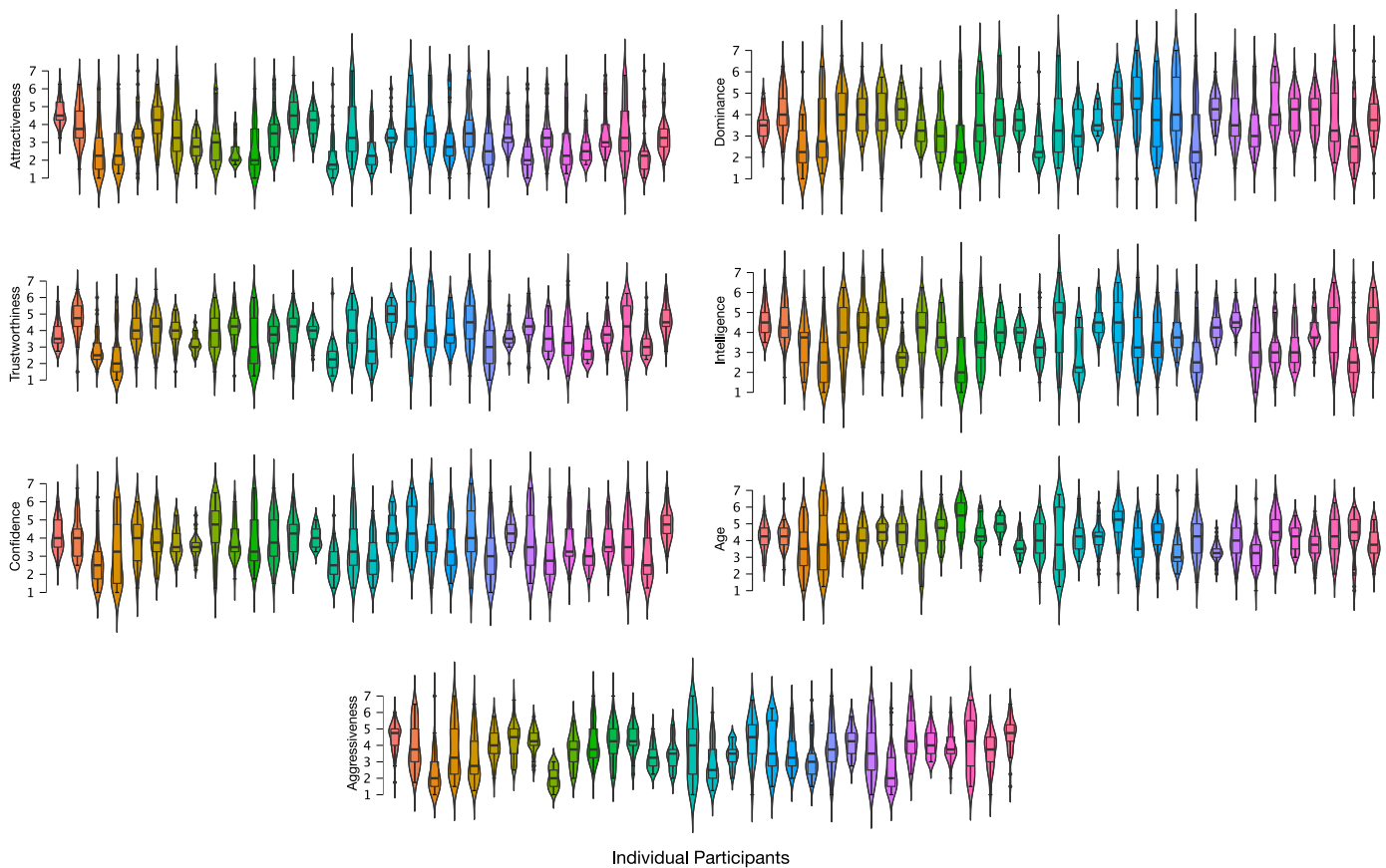


Figure 2. The summary of all 32 participants ratings on seven social characteristics. Each color represents one individual participant, the same color represented the same participant at different subplots. In each figure, the raw ratings of each participants were presented by a violin shape (showing a general shape of the rating distribution) as well as a boxplot element (showing the details of the rating distribution with third quartile, median, and the first quartile).

$p = 0.003$, $\eta^2_p = 0.12$). However, as suggested by the post-hoc analysis with Bonferroni corrections, such difference is solely driven by a lower level of variance at *Age* ($p = 0.114$ against *Attractiveness*; $p = 0.003$ against *Confidence*; $p = 0.006$ against *Dominance*; $p = 0.03$ against *Intelligence*; $p = 0.003$ against *Aggressiveness*). The differences among the other six characteristics did not reach a level of significance (all $ps > 0.95$). This is reasonable because the *Age* is different from the other six characteristics because it can be measured directly by years. In general, for at least the six characteristics,

it is likely that there is a similar degree of variability in the different ratings.

Derivative of Gaussian fitting

To quantify the serial dependence in seven facial characteristics, we pooled the judgment errors of all participants and fitted the DoG. It was found that judgment errors of every facial characteristic were positively related to the difference of facial trait

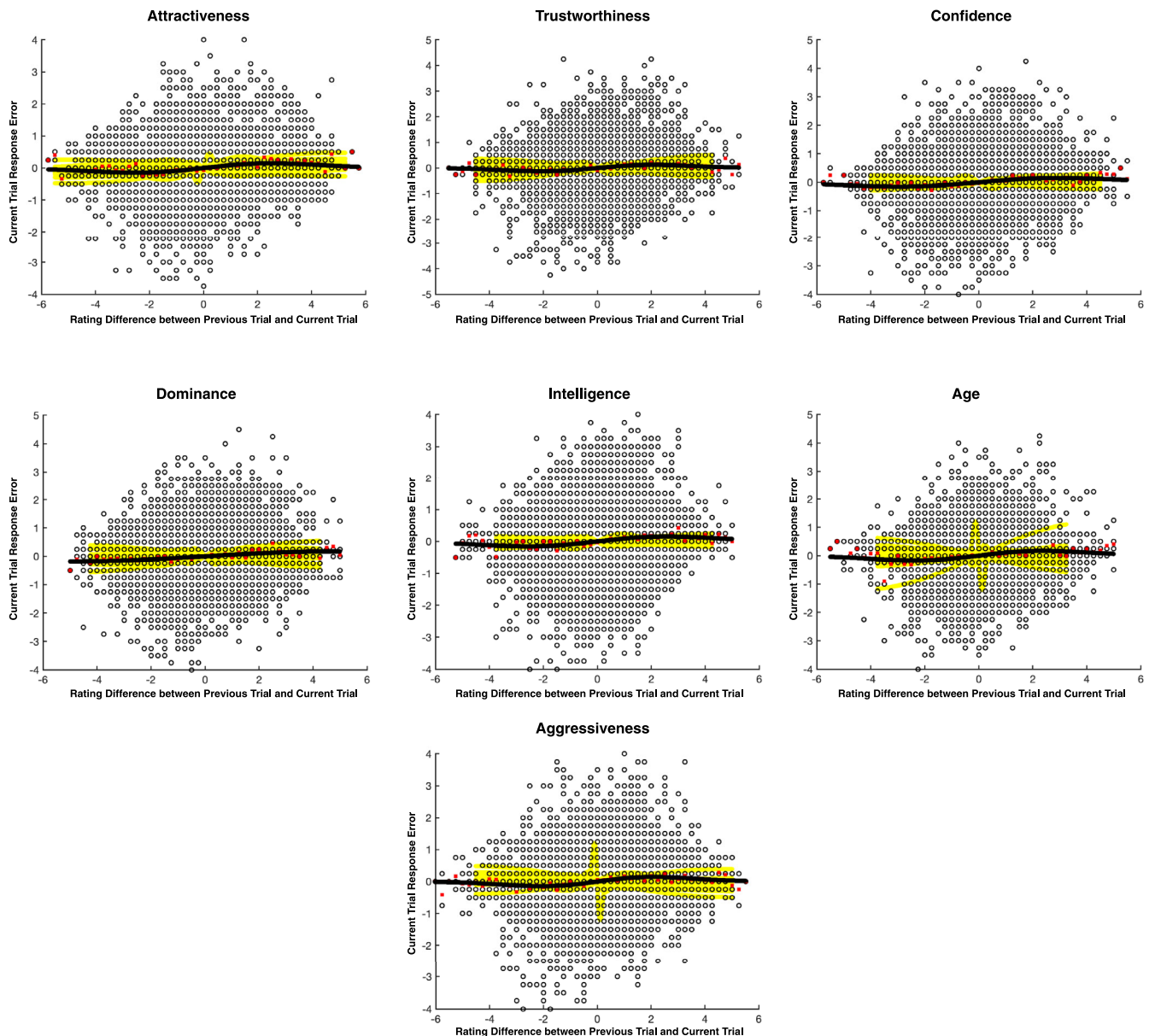


Figure 3. Serial dependence and DoG fittings for each facial trait. Y-axis indicated judgment error for the present face, and x-axis indicated differences in trait value between the previous and the present face. Each hollow-black dots represent the individual responses. Each solid-red dots represent the average of all participants. The bold black line is the fitted line of the DoG, and the yellow area represents the individual DoG fits.

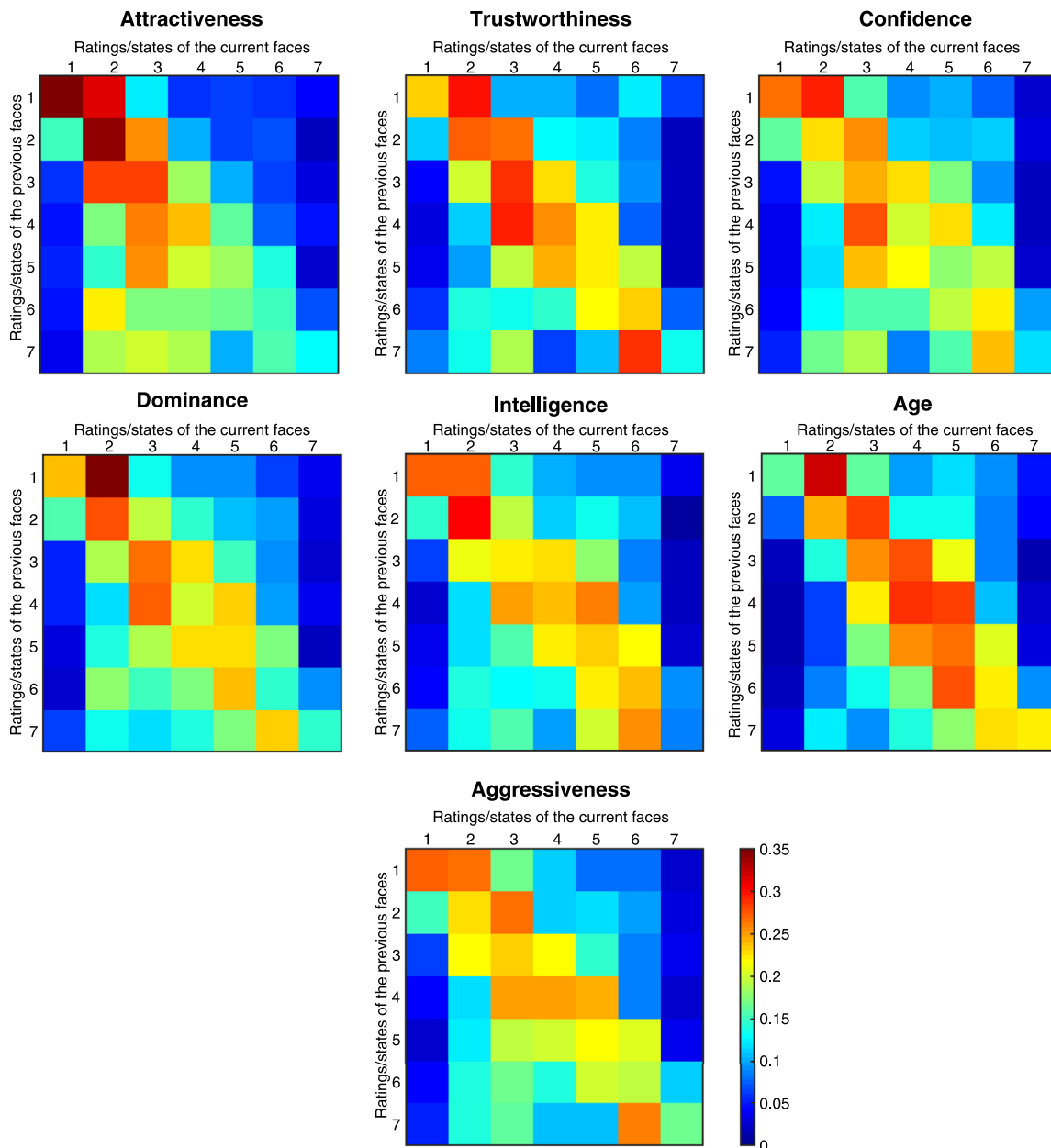


Figure 4. Transitional probability matrices of the seven facial traits. For each individual figure, the color of each cell represents the transitional probabilities of each current response (y-axis) given a previous response (x-axis). Higher transitional probabilities are marked as red, whereas lower transitional probabilities are marked as blue.

scores between the previous and present faces: as for attractiveness, participants would judge the present face as more attractiveness when the previous face was more attractive than the present one (Figure 3). The amplitude (a) of DoG function was calculated specifically for each facial characteristic as the size of this serial dependence effect (positive a indicates serial dependence occurs). We then compared the actual a of each function against the null distribution of a s by permutation. The results suggested that serial dependence (SD) occurs at all of the seven facial traits,

including attractiveness ($a = 0.15, p < 0.01$), confidence ($a = 0.16, p < 0.01$), dominance ($a = 0.15, p < 0.05$), intelligence ($a = 0.16, p < 0.01$), age ($a = 0.16, p < 0.01$); and marginally significant at trustworthiness ($a = 0.12, p = 0.090$), as well as aggressiveness ($a = 0.14, p = 0.075$).

To further investigate the pattern(s) of serial dependence among the perception of the seven facial characteristics, we compared the a s of each facial characteristic by a one-way repeated-measures analysis of variance. The results suggested that the main

effect of facial characteristics was not significant (Greenhouse-Geisser corrected $F[3.95, 122.40] = 1.20$, $p = 0.31$, $\eta^2_p = 0.037$), suggesting that the serial dependences between facial characteristics are highly similar, which supports the notion that there is a general pattern of serial dependence.

Markov Chain modeling

Alongside the traditional DoG fitting, we used the discrete-time Markov Chain modeling to qualify the serial dependence. We pooled the ratings of all the participants in separate facial characteristics and fitted them with the Markov Chain model. The results were represented on a 7×7 transitional probability matrices for each characteristic respectively (Figure 4). It is obvious that across all seven facial characteristics at all possible states, the rating of the current trial closely resembles that of the previous trial, indicating the existence of serial dependence for each facial characteristic. To qualify the existence of the serial dependence of these social evaluations of faces, we compared each transitional probability matrix against the diagonal matrix (which represents a pattern of strictly same rating after a given rating). The results of the matrix correlation analysis suggested that all matrices significantly resembled the diagonal matrix (all $r_s > 0.34$, all $p_s < 0.017$). Moreover, the information entropy ($H_{\text{mean}} = 2.55$) of each state at each facial trait is lower than that of a hypothetical state without serial dependence ($H_0 = \log_2(7) = 2.81$; at a hypothetical situation that all transitional probabilities are equal to $1/7$). Therefore, with the impact of the rating of the previous trial, the perception of the current trial is with less uncertainty (reflected by a lower entropy). Convergent evidence from DoG fitting and Markov Chain modeling together suggested that the serial dependence occurs at the social evaluation of faces.

To qualify the similarities among the serial dependence pattern(s), we also conducted the matrix correlation analysis to compare the similarities of the serial dependencies among the seven facial traits. The results indicated that the matrices of the seven facial characteristics were correlated significantly with each other (Table 1), which suggested that the serial dependence occurs at all facial traits and the serial dependence among facial characteristics is highly similar. Thus the perception of these facial traits may share a general pattern of serial dependence.

Further analysis based on Markov Chain modeling

Based on the theory of efficient coding, the precision of the inner representation at different stages is

| | Att | Tru | Con | Dom | Int | Age | Agg |
|-----|---------|---------|---------|---------|---------|---------|-----|
| Att | | | | | | | |
| Tru | 0.83*** | | | | | | |
| Con | 0.86*** | 0.94*** | | | | | |
| Dom | 0.87*** | 0.92*** | 0.93*** | | | | |
| Int | 0.81*** | 0.94*** | 0.94*** | 0.93*** | | | |
| Age | 0.68*** | 0.88*** | 0.86*** | 0.89*** | 0.87*** | | |
| Agg | 0.86*** | 0.95*** | 0.96*** | 0.93*** | 0.94*** | 0.89*** | |

Table 1. Correlation Coefficients Among the Seven Transitional Probability Matrices. Notes: Att, attractiveness; Tru, trustworthiness; Con, confidence; Dom, dominance; Int, intelligence; Agg, aggressiveness. *** $p < 0.001$; all p values are Bonferroni corrected.

proportional to “the frequency with which the state is encountered” (Polania, Woodford, & Ruff, 2019). Consequently, representations at the state/rating with higher probability shall be more precise (with higher SNR, Burr & Cicchini, 2014; Polania et al., 2019), which means that the perception of that state/rating is less relied on the “aid” of the serial dependence (Cicchini et al., 2018). If the perception at a state is less affected by the previous state (i.e., weaker serial dependence), then many other states can transfer to this state with evenly distributed chance, leaving high-information entropy. Therefore the information entropy of the transitional probability matrix at each stage reflects the strength of the transitional pattern at that state (a reflection of the strength of serial dependence). Therefore the similarity between the distribution of information entropy of the transitional probability matrix and the probability distribution of each rating level can be a good indicator of the efficient coding strategy. The distribution of each rating (red line) and the change of information entropy of the transitional probability matrices (blue line) at each social characteristic closely resembled each other (Figure 5). To test the similarity of distribution of rating and that of information entropy of transitional probability matrices, we conducted Kolmogorov-Smirnov testing and the results showed that the difference between the entropy distribution and the rating distribution was not significant for each facial trait (all $p_s > 0.20$), which indicated highly similar pattern between the two distributions at all seven facial characteristics. This finding suggested that the serial dependence found here follows the efficient coding strategy.

Discussion

In this study, we examined the pattern of serial dependence by testing participants with judging the

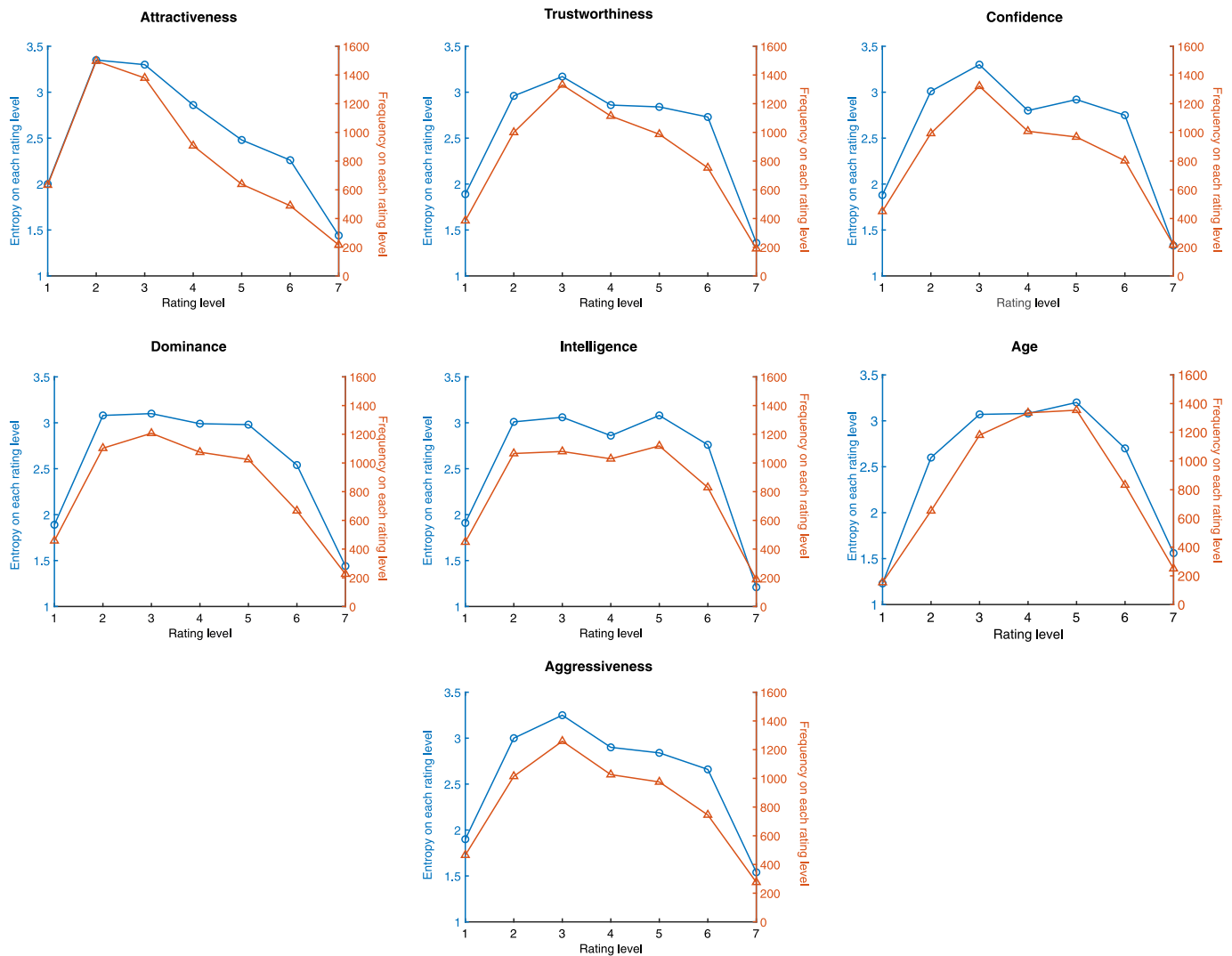


Figure 5. The distribution of each rating (red line) and the change of information entropy of the transitional probability matrices (blue line) at each rating state of each social characteristic. The change of the information entropy distributions closely resembles the distribution of the ratings. Again, this suggests the serial dependence (reflected by the distribution of the information entropy of the transitional probability matrices).

social facial characteristics of faces. By measuring and analyzing the serial dependence among seven facial traits using two statistical methods, this study provided new evidence suggesting that serial dependencies among facial traits are highly similar and temporal stability in judgment for different facial traits is achieved via a general pattern of serial dependence. The general pattern can be found at (1) the comparative magnitude of serial dependence; (2) the similar transitional pattern of serial dependence; and (3) the similar functional role of serial dependence. Using both DoG fitting and the Markov Chain modeling, we found that the serial dependence is highly consistent among seven dramatically different facial characteristics. Further analysis of the transitional probability matrix (of the Markov Chain) also suggests that the serial dependence

of these seven characteristics has the same transitional pattern, which can forecast the perception of the future state. Moreover, further analysis on the information entropy of the transitional probability matrices (of Markov Chain analysis) suggested that the serial dependence follows the efficient coding strategy.

Apart from the conventional method of fitting data with the DoG, some previous studies also used the Kalman filter model to quantify the serial dependence (Burr & Cicchini, 2014; Cicchini et al., 2018). The Kalman-filter model is an effective method to reveal the mechanism and function of serial dependence. It not only helped to explore how the serial dependence changed as the function of inter-stimulus difference by analyzing the weight change of the previous stimuli (Cicchini et al., 2018), but also showed that serial

dependence increases the integration of the successive sensory information, improving the efficiency of visual processing (Burr & Cicchini, 2014; Cicchini et al., 2018). Our study uses Markov Chain modeling as a new analysis method to measure the serial dependence. Markov Chain modeling serves as a good addition to DoG fitting and Kalman-filter modeling because it allows us to interpret the serial dependence in a totally different way than existing methods. As the visual system achieves temporal stability via serial dependence, by which perception of the current visual input is biased toward visual input from the recent past (Lieberman et al., 2014), the process of maintaining temporal stability can be considered as a kind of Markov process, during which the current perceptual states (in this case, the ratings) depend only on the previous states. The Markov chain modeling is able to demonstrate the transitional probability matrix among several perceptual states for each facial trait judgment. Using the Markov Chaining modeling, we can better observe how the ratings/states of the facial traits transitioned via the transitional probability matrix. Therefore the distribution of these transitional matrices directly revealed the patterns of serial dependence of facial trait judgments. Moreover, the finding from the information entropy analysis showed a similar distribution between the information entropy of transitional matrices and the (natural) distribution of each facial trait (based on the rating), which suggested that the serial dependence follows the efficient coding strategy. This finding provides further evidence to support the functional role of serial dependence. Future research may consider further studying the pattern within different transitional probability matrices.

In this study, we used the same stimuli and instructed the participants to rate the faces of different facial traits. Doing so allowed us to test serial dependence among different perceptual tasks while limiting the potential confounding factors when testing different stimuli. This manipulation has been used in related research areas like ensemble coding. Haberman and colleagues (2015) studied the similarity and difference of ensemble coding across different levels of visual processing. They used facial identity and facial expression to represent the high-level vision. Although they did not use the very same facial images, one would agree that the multifaceted nature of the face allows it to be a great stimulus to unveil the potential similarity and difference of a kind of perceptual phenomenon.

The role of serial dependence has been widely discussed. Consistent evidence has supported that serial dependence is an efficient mechanism in visual processing that “exploits the redundancy of the visual scene as an optimization strategy” (Cichini et al., 2018). Following this previous attempt, the results of the current study suggest that the efficient coding strategy drives serial dependence. Via serial dependence, our

brains adapt the efficiency of visual processing to the distribution of environmental stimuli, which means that our visual system is able to dynamically adjust the precision of coding based on the frequency at which we encounter the stimuli. This hypothesis is consistent with some previous observations. For example, it has been found that serial dependence was most pronounced for unfamiliar faces that were presented to participants with relatively low frequency and vice versa (Kok et al., 2017). Our results, in line with previous findings, indicate that serial dependence follows the efficient coding strategy, leading to more efficient perception: to achieve an optimal perceptual strategy (Cichini et al., 2018).

The general mechanism underlying serial dependence of face perception offers new evidence for a better understanding of social evaluation in face processing. Previous studies have shown that people make facial judgments in social context based on high-level properties of facial appearance (e.g., masculinity), as well as a multiple-channel representation of facial appearance at early age of visual processing, suggesting a task-specific mechanism in social trait evaluation (Balas & Verdugo, 2018). For example, evaluation of trustworthiness depends critically on horizontal orientation, whereas competence and dominance evaluation do not, which indicates that the visual information used for social evaluation relies on which trait the observers are trying to evaluate (Balas & Verdugo, 2018). Whereas previous studies emphasized the task-specific mechanisms in social trait judgment, our results, which showed highly similar serial dependence in different kinds of facial trait judgments, indicate that social evaluation in face processing also operates via a similar mechanism to maintain temporal stability. Our results support a common neural basis and mechanism underlying different social evaluations for faces. Future studies may expand this study by measuring the neural circuits of the serial dependence.

Recent studies have demonstrated that serial dependence not only originates from perceptual processing but also involves post-perceptual processing, for example, the decision stage (Ceylan et al., 2021; Fritsche et al., 2017). Judgment of different facial traits in a social context is related to different decision-making with regard to faces. For example, the two-dimensional model reduces the trait judgments into two underlying dimensions: trustworthiness/valence and dominance (Oosterhof & Todorov, 2008). The trustworthiness/valence dimension concerns intention to help or harm, whereas the dominance dimension relates to the perceived ability to carry out the helpful or harmful intention (Oosterhof & Todorov, 2008). Therefore it is reasonable to assume that the visual system achieves temporal stability for different facial trait judgments via different serial dependence mechanisms. However, our results turn against such

a hypothesis and supported a common and general mechanism underlying serial dependence across different visual tasks. Thus the current study further suggests that the serial dependence is an omnipresent aspect of visual processing that overrides individual visual tasks.

Inevitably, there are several shortcomings in our study. First, this study focused on social facial trait judgment, which is indeed an important aspect of high-level vision. However, to achieve a better understanding of serial dependence of vision, other stimuli shall be tested in the future. Second, participants here were asked to rate the faces on a seven-point Likert scale. This paradigm is less sensitive than the adjust-to-match methods used by some previous studies (e.g., Liberman et al., 2014); however, we must admit that using Likert scale allowed us to model the data via Markov Chain easier (fewer possible states allow a clearer output). Subsequent studies may consider using other paradigms.

In conclusion, even though recent studies have investigated visual serial dependence, it is still unknown whether there exists one general pattern of serial dependence or different patterns of serial dependencies for individual visual tasks. Also, we do not fully clear the pattern of bias in serial dependence. In this study, we tried to answer these questions by measuring serial dependence of several facial characteristics. Using both Markov Chain modeling and DoG fitting, we found that serial dependence occurs in face perception of social characteristics. Further investigation showed that it is highly likely that the perception of these facial characteristics shares a general pattern of serial dependence. Finally, further analysis of the information entropy of the transitional probability matrices (of Markov Chain analysis) suggested that serial dependence follows the efficient coding strategy. Thus these findings support the general serial dependence hypothesis and the notion that the high-level face perception is indeed hierarchical. Moreover, by introducing Markov Chain modeling in a serial dependence study, we offer a new analysis protocol to better investigate this perceptual phenomenon.

Keywords: serial dependence, face perception, social characteristics, Markov process

Acknowledgments

The authors thank Weiying Yang for the help of data collection.

H. Ying is supported by the Natural Science Foundation of Jiangsu Province (BK20200867), and the Entrepreneurship and Innovation Plan of Jiangsu

Province. J.M. Yu conducted this study under the Undergraduate Research Advising Project.

Commercial relationships: none.

Corresponding author: Haojiang Ying.

Email: hjying@suda.edu.cn.

Address: Department of Psychology, School of Education, Soochow University, Suzhou, China.

References

- Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, *14*(3), 257–262.
- Alais, D., Leung, J., & Van der Burg, E. (2017). Linear Summation of Repulsive and Attractive Serial Dependencies: Orientation and Motion Dependencies Sum in Motion Perception. *Journal of Neuroscience*, *37*(16), 4381–4390.
- Balas, B., & Verdugo, M. Q. (2018). Low-level orientation information for social evaluation in face images. *Psychonomic Bulletin & Review*, *25*(6), 2224–2230.
- Ballem, C. C., & Todorov, A. (2007). Predicting political elections from rapid and unreflective face judgments. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(46), 17948–17953.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436.
- Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*(3), 305–327.
- Bliss, D. P., Sun, J. J., & D’Esposito, M. (2017). Serial dependence is absent at the time of perception but increases in visual working memory. *Scientific Reports*, *7*(1), 14739.
- Burns, E. J., Yang, W., & Ying, H. (2021). Friend effects framework: Contrastive and hierarchical processing in cheerleader effects. *Cognition*, *212*, 104715.
- Burr, D., & Cicchini, G. M. (2014). Vision: efficient adaptive coding. *Current Biology*, *24*(22), 1096–1098.
- Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience*, *6*(8), 641–651.
- Ceylan, G., Herzog, M. H., & Pascucci, D. (2021). Serial dependence does not originate from low-level visual processing. *Cognition*, *212*, 104709.
- Chen, L. F., & Yen, Y. S. (2007). Taiwanese Facial Expression Image Database. *Brain Mapping Laboratory, Institute of Brain Science*. Taipei: National Yang-Ming University.

- Cicchini, G. M., Mikellidou, K., & Burr, D. C. (2018). The functional role of serial dependence. *Proceedings. Biological Sciences*, 285(1890), 20181722.
- Clifford, C., Watson, T. L., & White, D. (2018). Two sources of bias explain errors in facial age estimation. *Royal Society Open Science*, 5(10), 180841.
- Eimer, M. (2000). The face-specific N170 component reflects late stages in the structural encoding of faces. *Neuroreport*, 11(10), 2319–2324.
- Etcoff, N. L., Stock, S., Haley, L. E., Vickery, S. A., & House, D. M. (2011). Cosmetics as a feature of the extended human phenotype: modulation of the perception of biologically important facial signals. *PLoS One*, 6(10), e25656.
- Fischer, J., & Whitney, D. (2014). Serial dependence in visual perception. *Nature Neuroscience*, 17(5), 738–743.
- Fritsche, M., Mostert, P., & de Lange, F. P. (2017). Opposite Effects of Recent History on Perception and Decision. *Current Biology*, 27(4), 590–595.
- Fritsche, M., Spaak, E., & de Lange, F. P. (2020). A Bayesian and efficient observer model explains concurrent attractive and repulsive history biases in visual perception. *eLife*, 9, e55389.
- Haberman, J., Brady, T. F., & Alvarez, G. A. (2015). Individual differences in ensemble perception reveal multiple, independent levels of ensemble representation. *Journal of experimental psychology. General*, 144(2), 432–446, <https://doi.org/10.1037/xge0000053>.
- Hägglström, O. (2002). *Finite Markov chains and algorithmic applications* (London Mathematical Society Student Texts). Cambridge: Cambridge University Press, doi:10.1017/CBO9780511613586.
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, 4(6), 223–233.
- Jones, B. C., DeBruine, L. M., Flake, J. K., Liuzza, M. T., Antfolk, J., & Arinze, N. C., ... & Sirota, M. (2021). To which world regions does the valence–dominance model of social perception apply?. *Nature Human Behaviour*, 5(1), 159–169.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, 17(11), 4302–4311.
- Kok, R., Taubert, J., Van der Burg, E., Rhodes, G., & Alais, D. (2017). Face familiarity promotes stable identity recognition: Exploring face perception using serial dependence. *Royal Society Open Science*, 4(3), 160685, <https://doi.org/10.1098/rsos.160685>.
- Liberman, A., Fischer, J., & Whitney, D. (2014). Serial dependence in the perception of faces. *Current Biology*, 24(21), 2569–2574.
- Manassi, M., Liberman, A., Chaney, W., & Whitney, D. (2017). The perceived stability of scenes: serial dependence in ensemble representations. *Scientific Reports*, 7(1), 1971.
- O’Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D. M., & Dolan, R. J. (2003). Beauty in a smile: The role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia*, 41(2), 147–155, [https://doi.org/10.1016/s0028-3932\(02\)00145-8](https://doi.org/10.1016/s0028-3932(02)00145-8).
- Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the United States of America*, 105(32), 11087–11092.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10(4), 437–442.
- Polanía, R., Woodford, M., & Ruff, C. C. (2019). Efficient coding of subjective value. *Nature Neuroscience*, 22(1), 134–142.
- Rhodes, G., & Jeffery, L. (2006). Adaptive norm-based coding of facial identity. *Vision Research*, 46(18), 2977–2987.
- Sutherland, C. A., Oldmeadow, J. A., Santos, I. M., Towler, J., Michael Burt, D., & Young, A. W. (2013). Social inferences from faces: ambient images generate a three-dimensional model. *Cognition*, 127(1), 105–118.
- Taubert, J., Alais, D., & Burr, D. (2016). Different coding strategies for the perception of stable and changeable facial attributes. *Scientific Reports*, 6, 32239.
- Todorov, A., Gobbini, M. I., Evans, K. K., & Haxby, J. V. (2007). Spontaneous retrieval of affective person knowledge in face perception. *Neuropsychologia*, 45(1), 163–173.
- Van der Burg, E., Rhodes, G., & Alais, D. (2019). Positive sequential dependency for face attractiveness perception. *Journal of Vision*, 19(12), 6, <https://doi.org/10.1167/19.12.6>.
- Wang, Y., Yao, P., & Zhou, G. (2015). The influence of facial attractiveness and personality labels on men and women’s mate preference. *Acta Psychologica Sinica*, 47 (1), 108–118.
- Webster, M. A. (2015). Visual adaptation. *Annual Review of Vision Science*, 1, 547–567.
- Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods*, 42(3), 671–684.

- Xia, Y., Leib, A. Y., & Whitney, D. (2016). Serial dependence in the perception of attractiveness. *Journal of Vision*, 16(15), 28.
- Yap, W. J., Chan, E., & Christopoulos, G. I. (July 2016). Nanyang facial emotional expression [N-FEE] database - development and validation. *Poster presented at the 23rd Congress of the International Association for Cross-Cultural Psychology*, Nagoya, Japan.
- Ying, H., Burns, E., Choo, A. M., & Xu, H. (2020). Temporal and spatial ensemble statistics are formed by distinct mechanisms. *Cognition*, 195, 104128–104128.
- Ying, H., & Chen, Y. (2021, May). A Neural Network Approach to Subjective Human Face Perception Classification based on Social Characteristics. In *2021 IEEE 10th Data Driven Control and Learning Systems Conference (DDCLS)* (pp. 457–462). IEEE.
- Zhang, G. L., Li, A. S., Miao, C. G., He, X., Zhang, M., & Zhang, Y. (2018). A consumer-grade LCD monitor for precise visual stimulation. *Behavior Research Methods*, 50(4), 1496–1502.