Serial dependence in the perception of attractiveness

Ye Xia Department of Psychology, University of California, Berkeley, CA, USA

Allison Yamanashi Leib Department of Psychology, University of California, Berkeley, CA, USA

David Whitney Department of Psychology, University of California, Berkeley, CA, USA
Helen Wills Neuroscience Institute, University of California, Berkeley, CA, USA
Vision Science Group, University of California, Berkeley, CA, USA

The perception of attractiveness is essential for choices of food, object, and mate preference. Like perception of other visual features, perception of attractiveness is stable despite constant changes of image properties due to factors like occlusion, visual noise, and eye movements. Recent results demonstrate that perception of low-level stimulus features and even more complex attributes like human identity are biased towards recent percepts. This effect is often called serial dependence. Some recent studies have suggested that serial dependence also exists for perceived facial attractiveness, though there is also concern that the reported effects are due to response bias. Here we used an attractiveness-rating task to test the existence of serial dependence in perceived facial attractiveness. Our results demonstrate that perceived face attractiveness was pulled by the attractiveness level of facial images encountered up to 6 s prior. This effect was not due to response bias and did not rely on the previous motor response. This perceptual pull increased as the difference in attractiveness between previous and current stimuli increased. Our results reconcile previously conflicting findings and extend previous work, demonstrating that sequential dependence in perception operates across different levels of visual analysis, even at the highest levels of perceptual interpretation.

Introduction

Humans make aesthetic judgments all the time about the attractiveness or desirability of objects and scenes. Aesthetic judgments are not merely about judging works of art; they are constantly involved in our daily activity, influencing or determining our choices of food, object (Creusen & Schoormans, 2005), and mate preference (Rhodes, Simmons, & Peters, 2005).

Aesthetic judgments are based on perceptual processing (Arnheim, 1954; Livingstone & Hubel, 2002; Solso, 1996). These judgments, like other perceptual experiences, are thought to be relatively stable in spite of fluctuations in the raw visual input we receive due to factors like occlusion, visual noise, and eye movements. One mechanism that allows the visual system to achieve this stability is serial dependence. Recent results have revealed that the perception of visual features such as orientation (Fischer & Whitney, 2014), numerosity (Cicchini, Anobile, & Burr, 2014), and facial identity (Liberman, Fischer, & Whitney, 2014) are systematically assimilated toward visual input from the recent past. This perceptual pull has been distinguished from hysteresis in motor responses or decision processes, and has been shown to be tuned by the magnitude of the difference between previous and current visual inputs (Fischer & Whitney, 2014; Liberman, Fischer, & Whitney, 2014).

Is aesthetics perception similarly stable like feature perception? Some previous studies have suggested that the answer is yes. It has been shown that there is a positive correlation between observers' successive attractiveness ratings of facial images (Kondo, Takahashi, & Watanabe, 2012; Taubert, Van der Burg, & Alais, 2016). This suggests that there is an assimilative sequential dependence in attractiveness judgments.

Citation: Xia, Y., Leib, A. Y., & Whitney, D. (2016). Serial dependence in the perception of attractiveness. *Journal of Vision, 16*(15):28, 1–8, doi:10.1167/16.15.28.

doi: 10.1167/16.15.28 Received July 13, 2016; published December 22, 2016 ISSN 1534-7362

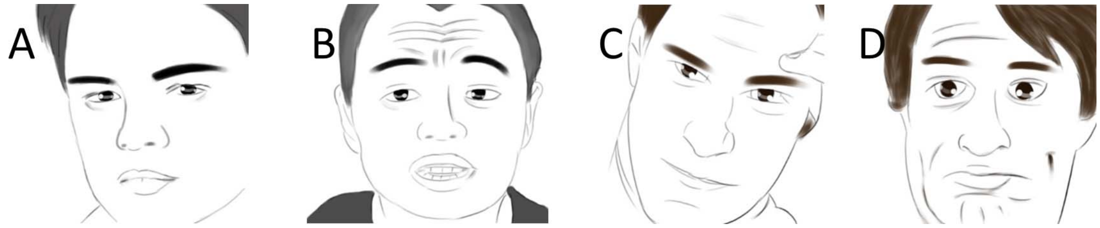

Figure 1. Cartoons of four examples of the experiment stimuli. (A–B) Two different face images of one identity; (C–D) two different face images of another identity. The stimuli of the experiment were 5,000 color photographs of 50 different identities. The four images in this figure are cartoon versions for reproduction.

However, Pegors, Mattar, Bryan, and Epstein (2015) have suggested that this assimilative sequential dependence is merely a response bias. They found that observers' ratings of facial attractiveness were attracted by ratings given to an orthogonal feature of previously seen faces (e.g., hair darkness) rather than the attractiveness of previous face images. There is, therefore, some debate about the existence of serial dependence in attractiveness perception.

Before we can conclude that there is (or is not) serial dependence in perceived attractiveness, there are some unanswered questions that need to be addressed. Does the assimilative sequential dependence in facial-attractiveness judgments still hold true after response bias is ruled out? Is the act of making a motor response necessary for revealing the effect? The goal here was to build on the prior work already described, and to test whether there is serial dependence in facial attractiveness, independent of action or response bias. The answers to these questions are basic but necessary for establishing the existence of sequential dependence in aesthetic judgments.

Experiment 1

Methods

Subjects

We recruited 191 subjects from Amazon Mechanical Turk. Fifty-six subjects were excluded because they had low test–retest reliability in their data (see later for exclusion criteria). Among the 135 remaining subjects, 69 were male, 59 were female, and seven did not report their gender. Subjects did not report their age.

Stimuli

We used 5,000 facial images from the FaceScrub data set (Ng & Winkler, 2014; Figure 1). These images belong to 50 different identities, 25 male and 25 female. Each identity has 100 face images (5,000 total images).

Procedure

The experiment was conducted online. There were two runs of 40 trials, with a pause in between. The pause duration was determined by the subject (self-paced). Each subject saw 40 different images of only one identity. The displayed identity was randomly chosen for each subject. The motivation for choosing only 40 images, rather than all 100, of the displayed identity was to make sure each run could be finished within about 5 min.

For each subject, the 40 chosen images were displayed in a random order in the first run, and then the same 40 images were displayed in another random order in the second run (Figure 2). Each face image was displayed at the center of the screen for 1 s. Subjects were asked to give a rating from 1 to 7 to indicate how attractive they perceived the face to be, with 7 being the highest attractiveness. Subjects responded either by clicking the corresponding button on the screen or pressing a corresponding key on the keyboard. The responses were self-paced. The web page automatically advanced to the next trial after each response. The intertrial interval depended on the subject's internet speed and typically varied between less than 0.5 s to 2 s.

Subject exclusion criteria

Since each subject rated the same images in two runs, any subject's responses in the two runs should be highly correlated unless the subject was giving either the same response for nearly all the stimuli or random responses independent of the stimuli. Permutation tests were conducted to find significance levels for correlations in each subject's ratings of the same stimuli in separate runs. The null hypothesis of the permutation tests was that the ratings were not specific to the facial images, and so any correlation between the ratings given in the two runs was due to chance. In order to get the distribution of the correlation under the null hypothesis, the image labels of the ratings of the first run were permuted 1,000 times. After each permutation, the correlation between the ratings of the two runs was calculated based on the permuted image labels and stored as one simulated correlation. The histogram of

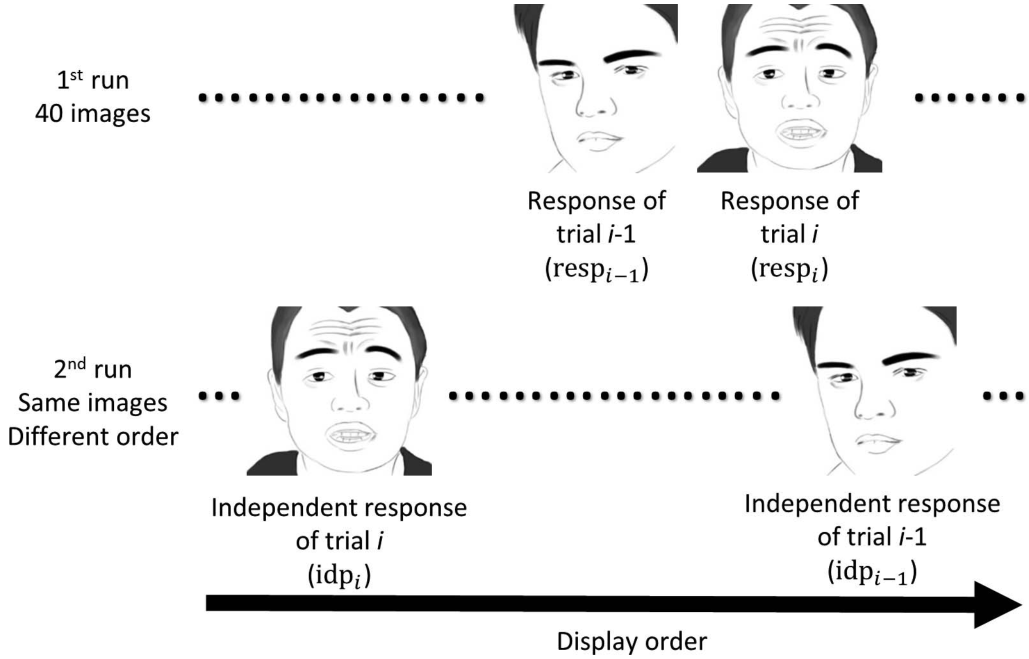

Figure 2. Illustration of experiment design and the terminology. Forty face images were rated in a random order in the first run and then the same 40 images were rated in another random order in the second run. In the analysis of the responses given in one run, the responses given in the other run can serve as independent references. For example, an independent reference for the response to the image displayed in the *i*th trial in the first run (resp_i) can be the response to the same image in the second run, which is defined as the independent response of the *i*th trial (idp_i). Note that this independent response of the *i*th trial was not necessarily given at the *i*th trial in the second run. Also, independent responses of trials *i* and $i - 1$ were not necessarily given in successive trials. The corresponding mathematical symbols used in the models are noted in parentheses.

the 1,000 simulated correlations approximates the distribution of the correlation under the null hypothesis. The significance level of the original correlation was calculated using this null distribution in a two-tailed manner. In order to correct for multiple testing, the false-discovery rates were calculated for each subject. A false discovery rate of 0.05 was used as the threshold. Fifty-six subjects were excluded according to this threshold. This high exclusion rate was likely due to the low motivation of online subjects, the absence of supervision from experiments, and fake responses given by automated software (Downs, Holbrook, Sheng, & Cranor, 2010; McCreadie, Macdonald, & Ounis, 2010; Paolacci, Chandler, & Ipeirotis, 2010). Our exclusion rate was lower than the one reported by McCreadie et al. (2010).

Results and discussion

Serial dependence of perceived facial attractiveness

One intuitive way of testing serial dependence of perceived facial attractiveness is to calculate the correlation between the attractiveness ratings given to present stimuli and previous stimuli. If perceived facial attractiveness is attracted by the attractiveness level of the previous facial image, there should be a positive correlation between current and previous responses. However, this positive correlation can also be induced by a response bias shared by successive trials, which is an important confounder (Pegors et al., 2015).

We developed another method to test the serial dependence of perceived facial attractiveness without being confounded by response bias. For a given trial, the perceived attractiveness of the present facial image was potentially biased by the previous facial image. In order to quantify this potential bias, an estimate of the attractiveness of the present stimulus is needed, independent of any influence by the particular preceding stimulus. Recall that each subject rated each facial image twice in two separate runs with different image-display orders. Therefore, for a trial in a given run, the response given by the same subject to the same facial image in the other run could serve as the independent estimate of perceived attractiveness. To clarify: For a trial in a given run, the trial in the other run that shared the same facial image is called its corresponding trial in the following; the response given at this corresponding trial is referred to as the independent response (Figure 2). Note that the independent response of the *i*th trial of a given run refers to the response given at the *corresponding* trial, not the response given to the *i*th position in the sequence (Figure 2, bottom row).

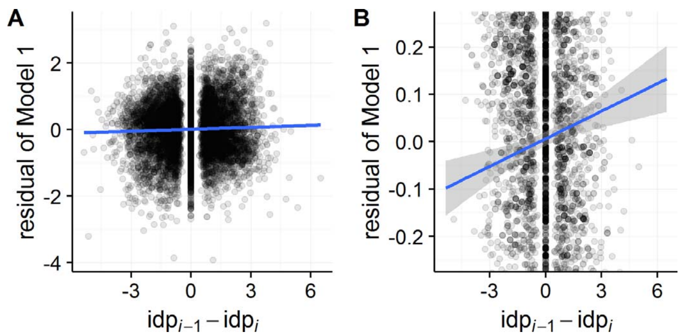

Figure 3. (A) Residuals of Model 1 plotted against the difference between the independent responses of previous and current trials. The blue line shows the linear regression and the gray shading represents the 95% confidence interval of the linear fitting. (B) Partial view of (A) with magnified y-axis for illustration of the confidence interval. The positive slope in the fit indicates that the response on the current trial was pulled toward the attractiveness of the previously seen face.

A simple linear model that predicts the response of a given trial by the independent response of that trial was fitted (Model 1), since the residuals of this model illustrate whether there is serial dependence for facial attractiveness. Model 1 can be formalized as

$$resp_i = \beta_0 + \beta_1 \cdot idp_i, \quad (1)$$

where $resp_i$ and idp_i are the response and independent response of the *i*th trial, respectively, and β_0 and β_1 are the coefficients of the model. Data from all the subjects were used collectively to fit the model. Raw ratings were z-scored beforehand within each subject and within each run.

The residuals of this model capture the variance in the responses that cannot be explained by the independent responses. If there is no serial dependence, the unexplained variance should be independent of the difference in attractiveness ratings between previous and current stimuli. The attractiveness difference between previous and current stimuli was measured by the difference between the independent responses of the previous and current trials. Note that this measurement is independent of the response bias of the current trial, because the independent responses were given in the independent run. Also note that the independent responses of previous and current trial were not necessarily taken from two successive trials in the independent run (Figure 2, bottom row).

The residuals of Model 1 were then plotted against the independent-response differences between previous and current stimuli (Figure 3). A linear regression indicated a significantly positive slope (slope = 0.017, *t*-test $P = 0.002$). This significantly positive slope suggests that current response depended not only on the response given to the same image by the same subject but also on the independent-response difference between previous and current trials. In order to quantify this influence, a more complicated model (Model 2) was fitted:

$$resp_i = \beta_0 + \beta_1 \cdot idp_i + \beta_2 \cdot (idp_{i-1} - idp_i), \quad (2)$$

where $resp_i$ is the response of the *i*th trial; idp_i and idp_{i-1} are the independent responses of trials i and $i-1$, respectively; and β_0, β_1, and β_2 are the coefficients of the model. The estimated value of β_2, 0.042, was defined as the effect size.

A permutation test was conducted to test the significance of β_2. The null hypothesis of the permutation test was that there was no real serial dependence and that any nonzero value of β_2 was due not to the sequence of the stimuli but to chance. In order to get the distribution of β_2 under the null hypothesis, the independent responses of one-back trials (idp_{i-1}) were permuted 1,000 times within each subject and each run (i.e., the permuted independent response of the one-back trial was actually the independent response of one random trial of the same run given by the same subject). After each permutation, Model 2 was fitted with the permuted previous independent responses and the value of β_2 was stored as one simulated coefficient value. The histogram of the 1,000 simulated coefficient values approximates the distribution of β_2 under the null hypothesis. The significance level of the original β_2 was calculated using this null distribution in a two-tailed manner. This permutation test confirmed that β_2 was significant ($P < 0.001$).

The fact that β_2 was significantly positive suggested that the response to the current image was pulled toward the attractiveness level of the previous stimulus. Note that the predicted variable (i.e., the response of the *i*th trial) and the independent responses used in the second regressor ($idp_{i-1} - idp_i$) were calculated from two different runs, so a positive value of β_2 could not be explained by response bias.

In order to further assure that the positive value of β_2 was not an artifact induced by our data-analysis method, we applied the same data-analysis method to test the influence of the stimulus of the next trial on the response of the current trial. This influence from the future cannot exist in reality; however, if our reported perceptual pull from the previous stimulus was merely an artifact due to the data-analysis method, a similar pull from the next stimulus would appear when the same data-analysis method was applied. Therefore, we replaced idp_{i-1} in Model 2 by idp_{i+1} to make Model 3:

$$resp_i = \alpha_0 + \alpha_1 \cdot idp_i + \alpha_2 \cdot (idp_{i+1} - idp_i), \quad (3)$$

where $resp_i$ is the response of the *i*th trial; idp_i and idp_{i+1} are the independent responses of trials i and $i+1$, respectively; and α_0, α_1, and α_2 are the coefficients of the model. The estimated value of α_2 was 0.010. The same permutation test showed that α_2 was not significantly different from zero ($P = 0.19$). The 95% bootstrapped

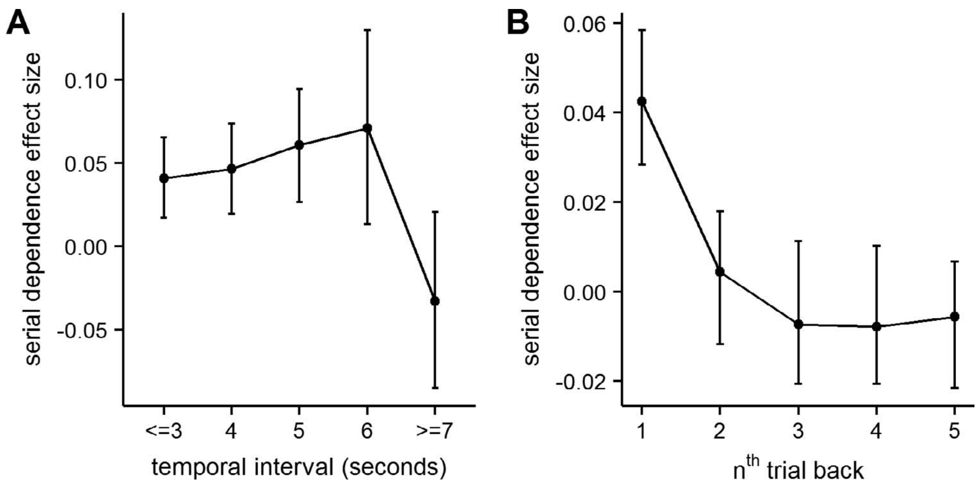

Figure 4. (A) Serial-dependence effect size as a function of the temporal interval between the stimulus displays of previous and current trials. (B) Serial dependence computed for stimuli presented one to five trials back from the present trial. Error bars in both panels represent the 95% bootstrapped confidence interval.

confidence interval of α_2 was [–0.005, 0.025], while the 95% bootstrapped confidence interval of β_2 was [0.027, 0.058]. Therefore, there was no artificial influence from the future, and this control simulation assured that our reported perceptual pull from the previous stimulus was not an artifact due to our data-analysis method.

Still significant with up to a 6-s delay

Because the responses were self-paced and subjects had different internet speeds, there were variations in the temporal intervals between the stimulus displays of successive trials. The typical range was between 3 and 7 s. The data were clustered into five groups according to the temporal interval between current trial and previous trial: less than or equal to 3 s, 4 s, 5 s, 6 s, and equal to or longer than 7 s. Model 2 was fitted to those groups separately. The 95% confidence intervals were calculated by bootstrapping. The result shows that the serial-dependence effect was significant for up to 6 s. However, after 6 s the data no longer showed a significant serial-dependence effect (Figure 4A). This indicates that previous image stimuli would significantly influence the perception of current image stimuli on the interval of 0–6 s. We also found that there was no significant influence from images two back or further from the current image (Figure 4B). It was only the immediately previous image that showed any kind of perceptual influence on the current stimuli.

Experiment 2

Experiment 1 showed that attractiveness ratings of facial images were assimilated toward the attractiveness of the previous facial images. A common concern in studies of sequential effects is that they might arise because of motor bias or sequential dependencies in the motor responses themselves (Shaffer, 1978; Wing & Kristofferson, 1973). The goal of this second experiment was to test whether the perceptual assimilation in Experiment 1 still holds true when no motor response is given in the previous trial.

Methods

Subjects

We recruited 256 subjects. Exclusion criteria were established using the same permutation-test process as in Experiment 1. A false-discovery rate of 0.05 was used as the threshold. We excluded 145 subjects according to this threshold. The causes of this high exclusion rate are described under Experiment 1. Subjects did not report their gender or age.

Stimuli

Stimuli were identical to those in Experiment 1.

Procedure

Experiment 2 followed most of the same design principles of Experiment 1. The following changes were made in order to test serial dependence in the absence of motor responses on the previous trial. The 40 face images of the displayed identity were randomly categorized into Groups A and B (20 images each). All the images of Group B were rated in both runs. One half of the images of Group A were randomly selected to be rated in the first run and not in the second run (when one image was not rated, the question page was not shown after the display of the image, so the subject could not give any response). The remaining half of the images from Group A were assigned to receive a rating in the second run but not in the first run. Each run consisted of trials alternating from a Group A to a Group B face image. Therefore, an image from Group B might be preceded by an image from Group A where no rating was given. Subjects could only respond using a mouse click. The motivation for this design was to make sure that subjects could not give a motor response when the question page was not displayed.

Results and discussion

We fitted Model 2 to the subset of trials with no response on the one-back trial. The estimated value of β_2 was 0.082. Note that for this experiment the value of β_2 should not be compared with 0 because there was a bias induced by the experiment design. In the subset of trials with no response on the one-back trial, the current trial stimulus always belonged to Group B and the one-back stimulus always came from Group A. Therefore, even if serial dependence was absent, there would be a negative correlation between ratings of current and previous trials, which would drive β_2 to be positive. The real baseline of β_2—i.e., the value of β_2 if there was no serial dependence—could be calculated by taking the mean of the β_2 values fitted from the permutation-test simulations. We then defined the effect size of serial dependence as the difference between the β_2 value fitted from real data and the baseline value of β_2. The effect size was equal to 0.064 and was significant according to the permutation test ($P < 0.001$). These results suggest that serial dependence of facial attractiveness did not depend on previous motor responses.

General discussion and conclusions

Our experiments demonstrated that perceived face attractiveness was pulled by the attractiveness level of facial images encountered up to 6 s ago. This effect was not due to response bias and did not rely on the previous motor response. This perceptual pull increased with increasing attractiveness difference between previous and current stimuli.

Previous studies have suggested a positive correlation between subjects' attractiveness ratings of the presented facial image and the previous facial image (Kondo et al., 2012, 2013; Kramer, Jones, & Sharma, 2013). This positive correlation could be the result of sequentially dependent attractiveness perception or sequentially dependent response bias (Pegors et al., 2015). For example, if a subject's rating criterion gradually changes over time (e.g., the subject tends to give higher ratings in the beginning of the experiment and lower ratings at the end of the experiment), then the autocorrelation of the subject's rating criteria will lead to a positive correlation between current and previous ratings. In contrast, we regressed attractiveness ratings on the difference of independent ratings of successive stimuli. The serial dependence we find is not accounted for by response biases.

Pegors et al. (2015) demonstrated a contrastive bias in the judgment of facial attractiveness. They showed subjects a sequence of faces, where each face was displayed for 4 s, and subjects were asked to make judgments about either the facial attractiveness or hair darkness in an alternating fashion. They found that attractiveness ratings shifted away from the mean attractiveness of the previous face. Although this finding appears to contradict the findings presented here, we do not believe that they preclude one another. Rather, due to the differences in stimulus-presentation duration, we believe that the findings are representative of two distinct visual mechanisms: a negative aftereffect from prolonged stimulus adaptation on the one hand, and serial dependence from shorter presentation time on the other. Which effect dominates depends on the duration of adaptation, as well as the time course of the positive and negative serial dependencies themselves. Indeed, negative aftereffects in face attractiveness have been reported before (Rhodes, Jeffery, Watson, Clifford, & Nakayama, 2003) and would naturally oppose any positive serial dependencies in perceived attractiveness. As in other domains, such as orientation (Fischer & Whitney, 2014) and face emotion (Taubert, Alais, & Burr, 2016), the time course of positive and negative aftereffects might be different.

Taubert, Van der Burg, and Alais (2016) demonstrated an assimilative perceptual bias of facial-attractiveness judgment using a two-alternative forced-choice paradigm. In their experiment, subjects gave a binary response (attractive or unattractive) for each face after a display of 300 ms. Their results showed that faces were more likely to be rated as attractive when the preceding face was attractive than when it was unattractive. Our results affirm theirs and extend them in three ways. First, the hysteresis of motor responses may also induce an apparent assimilative bias in attractiveness judgment. Experiment 2 shows that our results are not due to the action or the response per se. Second, we studied the serial dependence of facial attractiveness with various intertrial intervals and showed significant serial dependence with delays up to 6 s after the previous image. Third, our stimulus-display time was 1 s, which reduces the uncertainty about the attractiveness of the stimulus image. Serial dependence of facial attractiveness exists even when subjects are more certain about their perception.

Serial dependence of face attractiveness shares some properties with serial dependence reported in other domains. Serial dependence has been reported in the perception of orientation (Fischer & Whitney, 2014), position (Liberman, Kosovicheva, & Whitney, 2014), facial recognition (Liberman, Fischer, & Whitney, 2014; Taubert, Alais, & Burr, 2016), facial emotion (Liberman & Whitney, 2016), numerosity (Cicchini et al., 2014), and more. For these feature domains, the assimilative bias follows a pattern where there is first an increase and then a reduction back to zero with increasing feature dissimilarity. Our results showed that serial dependence of perceived facial attractiveness also increased with increasing difference in attractiveness between successive stimuli. The absence of the trend back to zero may be accounted for by the limited attractiveness differences between facial images of one displayed identity. The temporal tuning is somewhat shorter here than in some previous studies (Fischer & Whitney, 2014; Liberman, Fischer, & Whitney, 2014). For example, Liberman, Fischer, and Whitney (2014) showed that the perception of facial identity is pulled toward the one-back stimulus seen on average ~7500 ms prior to the current trial face. Our results revealed serial dependence of facial attractiveness for stimuli seen up to 6 s prior. However, the tasks and timing differed between the current study and the previous ones, so comparisons should be made with caution. Taking the same example, Liberman, Fischer, and Whitney (2014) used an adjustment task with a longer average delay between stimulus and response, whereas we used a scalar rating task right after the stimulus face. The density of trials per unit time was therefore somewhat higher in our experiment, and serial dependence in perception of different attributes might depend on delay as well as the number of intervening trials or stimuli (Fischer & Whitney, 2014). Further, an important consideration is that different features and object properties may have different autocorrelations in the physical world. For example, the attractiveness of a single face can change from moment to moment (Post, Haberman, Iwaki, & Whitney, 2012) but identity may be more stable. Therefore, if sequentially dependent perception mirrors the physical world, there may be differences in its temporal tuning (Taubert, Alais, & Burr, 2016).

Previous studies have posited the presence of a continuity-field mechanism that could account for serial dependence (Fischer & Whitney, 2014), shaping apparent stability by introducing sequential dependence to perceptual representations. The results here could support the existence of a continuity field that operates on aesthetic judgments like face attractiveness. Whether there are different continuity fields for different objects and features remains unclear. Furthermore, it remains unclear whether the continuity field or the serial dependence here will extend to other aesthetic judgments, such as those regarding paintings, architecture, nature, and so on. Similar to previous reports of serial dependence in other domains like orientation (Fischer & Whitney, 2014), the serial dependence of facial attractiveness seems to reduce sensitivity. But that is only true for random—unnatural—sequences of stimuli. If stimuli in the real world are physically autocorrelated, then serial dependence could reflect this and in that way be adaptive and possibly even predictive. Aesthetic judgments such as the perception of facial attractiveness are therefore more stable than would otherwise occur.

Keywords: aesthetics, facial attractiveness, serial dependence, stability

Acknowledgments

We thank N. Rodriguez and A. Nguyen for programming assistance; A. Liberman and W. Chaney for discussions and comments; J. Gagnon-Bartsch for statistics consultation; Z. Ryan, M. Fain, and K. Zhang for edits; and two referees for constructive comments. This work was supported by NIH Grant EY018216 (DW). Publication was made possible in part by support from the Berkeley Research Impact Initiative sponsored by the UC Berkeley Library.

Commercial relationships: none.
Corresponding author: Ye Xia.
Email: yexia@berkeley.edu.
Address: Department of Psychology, University of California, Berkeley, CA, USA.

References

Arnheim, R. (1954). *Art and visual perception: A psychology of the creative eye*. Berkeley, CA: University of California Press.

Cicchini, G. M., Anobile, G., & Burr, D. C. (2014). Compressive mapping of number to space reflects dynamic encoding mechanisms, not static logarithmic transform. *Proceedings of the National Academy of Sciences, USA*, *111*(21), 7867–7872.

Creusen, M. E., & Schoormans, J. P. (2005). The different roles of product appearance in consumer choice. *Journal of Product Innovation Management*, *22*(1), 63–81.

Downs, J. S., Holbrook, M. B., Sheng, S., & Cranor, L.

F. (2010, April). Are your participants gaming the system?: Screening Mechanical Turk workers. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2010,* 2399–2402.

Fischer, J., & Whitney, D. (2014). Serial dependence in visual perception. *Nature Neuroscience, 17*(5), 738–747.

Kondo, A., Takahashi, K., & Watanabe, K. (2012). Sequential effects in face-attractiveness judgment. *Perception*, *41*(1), 43–49.

Kondo, A., Takahashi, K., & Watanabe, K. (2013). Influence of gender membership on sequential decisions of face attractiveness. *Attention, Perception, & Psychophysics*, *75*(7), 1347–1352.

Kramer, R. S., Jones, A. L., & Sharma, D. (2013). Sequential effects in judgments of attractiveness: The influences of face race and sex. *PLoS One*, *8*(12), e82226.

Liberman, A., Fischer, J., & Whitney, D. (2014). Serial dependence in the perception of faces. *Current Biology*, *24*(21), 2569–2574.

Liberman, A., Kosovicheva, A., & Whitney, D. (2014). Serial dependence of position perception. *Journal of Vision*, *14*(10): 1186, doi:10.1167/14.10.1186. [Abstract]

Liberman, A., & Whitney, D. (2016). The serial dependence of object perception is independent of decision. *Journal of Vision*, *16*(12): 1313, doi:10.1167/16.12.1313. [Abstract]

Livingstone, M., & Hubel, D. H. (2002). *Vision and art: The biology of seeing* (Vol. 2). New York: Harry N. Abrams.

McCreadie, R. M., Macdonald, C., & Ounis, I. (2010, July). Crowdsourcing a news query classification dataset. *Proceedings of the ACM SIGIR 2010 Workshop on Crowdsourcing for Search Evaluation (CSE 2010), 2010*, 31–38.

Ng, H.-W., & Winkler, S. (2014, October). *A data-driven approach to cleaning large face datasets. 2014 IEEE International Conference on Image Processing (ICIP), 2014*, 343–347.

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, *5*(5), 411–419.

Pegors, T. K., Mattar, M. G., Bryan, P. B., & Epstein, R. A. (2015). Simultaneous perceptual and response biases on sequential face attractiveness judgments. *Journal of Experimental Psychology: General*, *144*(3), 664–673.

Post, R., Haberman, J., Iwaki, L., & Whitney, D. (2012). The frozen face effect: Why static photographs may not do you justice. *Frontiers in Psychology, 3,* 22.

Rhodes, G., Jeffery, L., Watson, T. L., Clifford, C. W., & Nakayama, K. (2003). Fitting the mind to the world: Face adaptation and attractiveness aftereffects. *Psychological Science*, *14*(6), 558–566.

Rhodes, G., Simmons, L. W., & Peters, M. (2005). Attractiveness and sexual behavior: Does attractiveness enhance mating success? *Evolution and Human Behavior*, *26*(2), 186–201.

Shaffer, L. (1978). Timing in the motor programming of typing. *The Quarterly Journal of Experimental Psychology*, *30*(2), 333–345.

Solso, R. L. (1996). *Cognition and the visual arts.* Cambridge, MA: MIT Press.

Taubert, J., Alais, D., & Burr, D. (2016). Different coding strategies for the perception of stable and changeable facial attributes. *Scientific Reports*, *6*, 32239.

Taubert, J., Van der Burg, E., & Alais, D. (2016). Love at second sight: Sequential dependence of facial attractiveness in an on-line dating paradigm. *Scientific Reports, 6,* 22740.

Wing, A. M., & Kristofferson, A. B. (1973). Response delays and the timing of discrete motor responses. *Perception & Psychophysics*, *14*(1), 5–12.