Positive sequential dependency for face attractiveness perception

Erik Van der Burg
Department Experimental and Applied Psychology, Vrije Universiteit, Amsterdam, The Netherlands
The School of Psychology, The University of Sydney, Sydney, NSW, Australia

Gillian Rhodes
The School of Psychological Science, The University of Western Australia, Perth, WA, Australia

David Alais
The School of Psychology, The University of Sydney, Sydney, NSW, Australia

Recent findings from several groups have demonstrated that visual perception at a given moment can be biased toward what was recently seen. This is true both for basic visual attributes and for more complex representations, such as face identity, gender, or expression. This assimilation to the recent past is a positive serial dependency, similar to a temporal averaging process that capitalizes on short-term correlations in visual input to reduce noise and boost perceptual continuity. Here we examine serial dependencies in face perception using a simple attractiveness rating task and a rapid series of briefly presented face stimuli. In a series of three experiments, our results confirm a previous report that face attractiveness exhibits a positive serial dependency. This intertrial effect is not only determined by face attractiveness on the previous trial, but also depends on the faces shown up to five trials back. We examine the effect of stimulus presentation duration and find that stimuli as brief as 56 ms produce a significant positive dependency similar in magnitude to that produced by stimuli presented for 1,000 ms. We observed stronger positive dependencies between same-gender faces, and found a task dependency: Alternating gender discrimination trials with attractiveness rating trials produced no serial dependency. In sum, these findings show that a perception-stabilizing assimilation effect operates in face attractiveness perception that is task dependent and is acquired surprisingly quickly.

Introduction

The visual system has a remarkable ability to find stable signals from an often noisy input stream that fluctuates due to numerous factors including eye movements, blinks, shadows, occlusion, etcetera. These fluctuations can often be disregarded as noise because many aspects of the visual environment are quite stable and predictable, with analyses showing there are significant short-term correlations in visual input (Dong & Atick, 1995). Recent studies have indicated there are mechanisms that function to stabilize the perception of visual features over time through an operation that pools input over the recent past to better predict the current input state (Kiyonaga, Scimeca, & Whitney, 2017). An averaging process of this kind (termed an "association field") is a sensible exploitation of temporal continuity but it should inevitably exert an assimilative influence on perception whereby current stimuli are judged to be more similar to the preceding stimulus than is physically the case. Several studies have shown that this is true for basic visual attributes such as orientation (Fischer & Whitney, 2014) and motion (Alais, Leung, & Van der Burg, 2017), as well as for more complex representations including numerosity (Cicchini, Anobile, & Burr, 2014; Corbett, Fischer, & Whitney, 2011) and scene perception (Manassi, Liberman, Chaney, & Whitney, 2017).

An assimilation toward recent stimulus history has also been found for several aspects of face perception. A number of studies using sequences of briefly presented faces have shown that face perception on a given trial is not independent but is influenced by the preceding face image. This serial effect is generally found to be a positive (or "assimilative") dependency, consistent with an integrating association field that averages over the recent past to stabilize perception. This has been shown for face identity (Liberman,

Citation: Van der Burg, E., Rhodes, G., & Alais, D. (2019). Positive sequential dependency for face attractiveness perception. *Journal of Vision*, *19*(12):6, 1–16, https://doi.org/10.1167/19.12.6.

https://doi.org/10.1167/19.12.6 Received April 4, 2019; published October 17, 2019 ISSN 1534-7362 Copyright 2019 The Authors

Fischer, & Whitney, 2014), face gender (Taubert, Alais, & Burr, 2016), and very recently for face attractiveness (Kok, Taubert, Van der Burg, Rhodes, & Alais, 2017; Taubert, Van der Burg, & Alais, 2016; Xia, Yamanashi Leib, & Whitney, 2016), eye gaze (Alais, Kong, Palmer, & Clifford, 2018), and body shape (Alexi et al., 2018). These studies all show that current face perception is assimilated toward the immediately preceding stimulus—an "attractive" serial dependency. However, not all serial dependencies are positive. For instance, Taubert, Alais, and Burr (2016) demonstrated that face expression exhibited a repulsive rather than attractive dependency on the preceding face, an effect more consistent with traditional repulsive perceptual aftereffects occurring after prolonged exposure to an adaptor. Taubert, Alais, and Burr's negative effect for face expression is therefore similar to repulsive aftereffects that are seen after several seconds of exposure to adaptor faces (cf. subsecond presentations in typical serial dependency studies) and which produce robust negative aftereffects for face expression (Fox & Barton, 2007; Hsu & Young, 2004), identity (Leopold, O'Toole, Vetter, & Blanz, 2001; Rhodes & Jeffery, 2006), gender (Webster, Kaping, Mizokami, & Duhamel, 2004), and attractiveness (Rhodes, Jeffery, Watson, Clifford, & Nakayama, 2003). Although exposure durations down to 1 s have been reported to produce repulsive face adaptation for identity (Leopold, Rhodes, Muller, & Jefferey, 2005), and other studies using brief stimuli have also found repulsive effects for visual orientation (Alais et al., 2017), auditory frequency perception (Alais, Orchard-Mills, & Van der Burg, 2015), and audiovisual temporal order (Van der Burg, Alais, & Cass, 2013, 2015; Van der Burg & Goodbourn, 2015), even shorter face presentations (<400 ms) have been shown to produce attractive aftereffects (Oruç & Barton, 2010).

It is not exactly clear why serial dependencies are sometimes positive and at other times negative, although factors such as exposure duration and stimulus reliability are important. Regardless of sign, each serial effect affords a functional benefit. The advantage of a positive dependence is that it improves signal reliability by pooling current and previous stimuli, reducing noise and thus boosting signal strength. Perceptually, this manifests as a priming effect where current perception is biased toward the recent past. While this has the disadvantage of blurring differences between successive stimuli, recent work modeling psychophysical data shows that the strength of positive dependence is determined by the relative reliability (i.e., the inverse of variance) of successive stimuli (Cicchini et al., 2014). When the current stimulus has lower reliability than the previous one, the current stimulus is down-weighted and the previous one is given greater weight, and vice versa. In this way, a positive serial dependence is an optimal method of combining inputs (Cicchini & Burr, 2018; Cicchini, Mikellidou, & Burr, 2018), similar to Bayesian maximum likelihood models (Alais & Burr, 2004; Ernst & Banks, 2002).

The functional advantage of a negative serial dependence, such as traditional perceptual aftereffects, is that it enhances discrimination around the adapted stimulus, improving our ability to see small differences. Exposure duration is an important factor, with negative dependences usually arising from relatively prolonged exposure and positive dependences usually requiring very brief exposure duration or a degree of stimulus uncertainty (Cicchini et al., 2014; Suárez-Pinilla, Seth, & Roseboom, 2018). Adding a delay between stimulus offset and the response can increase the positive serial effect (Bliss, Sun, & D'Esposito, 2017; Fritsche, Mostert, & de Lange, 2017), with longer delays increasing the serial dependency—although there is still an initial serial effect that occurs early (Manassi, Liberman, Kosovicheva, Zhang, & Whitney, 2018). In some cases, two different attributes within the same stimulus can show opposite dependences simultaneously, as in face attractiveness and gender (Taubert, Alais, & Burr, 2016) and orientation and motion (Alais et al., 2017), and a change in task with the same stimuli can also change the sign of the dependence, as observed for audiovisual relative timing (Keane, Bland, Matthews, Carroll, & Wallis, in press; Roseboom, 2019).

As noted above, there are reports of both positive (Kok et al., 2017; Liberman et al., 2014; Oruç & Barton, 2010; Taubert, Van der Burg, & Alais, 2016; Xia et al., 2016) and negative (Leopold et al., 2005; Rhodes et al., 2003; Taubert, Alais, & Burr, 2016; Webster et al., 2004) serial dependencies in face perception. There are numerous possible reasons for these conflicting results, such as different exposure durations, different face attributes being adapted, and methodological differences in adaptation and response. The key differences may be perceptual or postperceptual and decisional. Indeed, effects at both levels could be involved and they may have different time courses, with attraction (Oruç & Barton, 2010) giving way to repulsion (Leopold et al., 2005) as adaptation time increases. In our recent study (Taubert, Van der Burg, & Alais, 2016), we obtained evidence that serial dependence for face attractiveness is perceptual in nature because the assimilative effects we observed when the faces on the current trial (t) and the preceding trial ($t - 1$) were upright were reduced when an inverted face preceded an upright face. As face-selective cells respond more vigorously to upright face stimuli (Taubert, Van Belle, Vanduffel, Rossion, & Vogels, 2015; Yovel & Kanwisher, 2005), this reduction suggests a sensory basis underlies the reduced intertrial effect and also argues against a decision-level account because the task and response remained the same regardless of face orientation. However, inverting faces between trials did not eliminate the effect entirely, indicating at least part of the face attractiveness serial dependency may be due to the perceptual decision. Here, we further examine how these two sources (i.e., perception and decision) may contribute to face attractiveness serial dependence by examining the time-course of the effect and whether it is task or gender dependent.

Most serial dependency studies have used quite brief stimulus presentation durations, commonly on the order of half a second or even less (Cicchini et al., 2014; Fischer & Whitney, 2014; Taubert, Van der Burg, & Alais, 2016), although some earlier studies used longer presentation times (Kondo, Takahashi, & Watanabe, 2012; Pegors, Mattar, Bryan, & Epstein, 2015). Here we investigate how brief the stimulus presentation can be and still produce a significant positive dependency. This is important as it may provide important information about the underlying mechanism. If the intertrial effects are observed for very brief intervals, then this may bolster the claim that face attractiveness effects are due to a perceptual process. Indeed, Olson and Marshuetz (2005) reported that participants are able to discriminate attractive faces from unattractive faces with an exposure of just 13 ms (see also Willis & Todorov, 2006). They therefore proposed that the ability to rate someone's attractiveness rapidly may be due to perceptual processes that influence decisions with little awareness or intention, and that this may result from low-level visual processing. We also examined whether the serial dependency is specific to the task by asking participants to alternate between a gender task and an attractiveness task over trials. If the serial dependency has a perceptual basis and if participants automatically process facial attractiveness features as recently proposed (see e.g., Ritchie, Palermo, & Rhodes, 2017), then simply viewing the face should be enough to evoke the effect, regardless of the task on the trial. Finally, we tested for gender-specific effects, reasoning that if a gender-specific attractiveness dependency is obtained, it is likely to be perceptual because the attractiveness rating task remains constant across trials but the dependency will have varied based on the gender specific features on the previous trial.

Experiment 1: Positive sequential dependency for face attractiveness

The aim of Experiment 1 was to replicate our previous finding that the perceived attractiveness of a given face in a sequence of briefly presented faces depends on the attractiveness of the previous face. If perceived attractiveness on a given trial (t) is independent of the randomly selected face presented in the preceding trial ($t - 1$), there will be no correlation between attractiveness ratings on consecutive trials. Previous studies have shown that serial dependencies exist for sequences of face stimuli and that they are generally positive (Kok et al., 2017; Liberman et al., 2014; Taubert, Van der Burg, & Alais, 2016; Xia et al., 2016), although face expression produces a negative serial dependency (Taubert, Alais, & Burr, 2016). Here we expect a positive or "assimilative" dependency for face attractiveness, consistent with Xia et al. (2016). Furthermore, we aim to investigate whether the intertrial effects are gender specific. If we find a gender-specific effect (that is, a positive serial dependence is observed only when the gender on trial t matches the gender on trial $t - 1$) then this would support the notion that the serial dependency is driven by a perceptual effect. Finally, we will analyze whether the observed sequential effects are predominantly driven by the face shown on the previous trial only, or whether faces from trials further back also have an influence on the current trial. If the brain stabilizes the perception of visual features (in our case, facial features) over time through an operation that pools input over the recent past (as proposed by Kiyonaga et al., 2017), then it is possible that faces from several trials back would influence the current trial percept if they fall within the temporal window. If we indeed find evidence that faces from several trials back affect the perceived attractiveness on the current trial, then it not only provides evidence for a pooling mechanisms over time, but also strengthens the notion that the intertrial effects are due to a perceptual effect, as it is hard to explain how all the responses from all those n-back trials could affect the current response.

Method

Participants

Based on our previous research in face serial dependence, we selected 16 students to participate in Experiment 1 (see also Taubert, Van der Burg, & Alais, 2016, experiment 1). All participants were naive as to the purpose of the experiment and were paid $AU20 per hour for their participation. One participant was excluded from the data analysis because of a lack of variation in the attractiveness ratings (85% of the faces were scored as not attractive), leaving 15 subjects' data for analyses. In this experiment and those that follow, participants gave informed consent, the experiments accorded with the Declaration of Helsinki, and all procedures were approved by the University of Sydney's Human Research Ethics Committee.

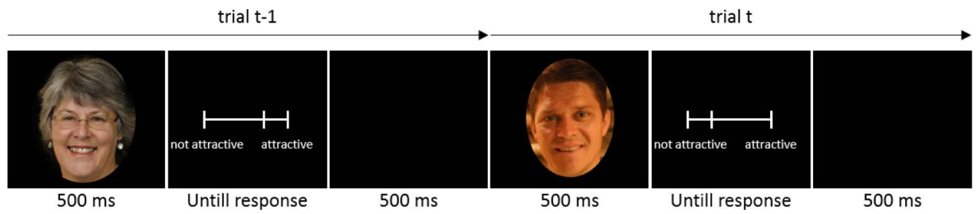

Figure 1. Illustration of the general paradigm used in the present study. Participants were shown a face for 500 ms that was randomly selected from a set of 242 faces, followed by a rating scale. The task was to rate the face's attractiveness using a mouse-controlled marker that moved continuously along a rating scale that spanned 201 ($\pm$100) pixels and whose endpoints were labeled "not attractive" (left end) and "attractive" (right end). A mouse-click recorded the marker's position on the 201-point attractiveness scale and initiated a blank screen for a 500-ms intertrial interval. The starting location of the adjustable marker was randomly determined on each trial by drawing from a flat distribution ranging from -100 to 100. The durations shown in this figure correspond to the timing in Experiment 1. Note that the faces shown in the figure are of authors DA and GR and were not used in the experiment. For privacy reasons, the faces used in the study cannot be shown.

Apparatus and stimuli

The experiment was programmed and run using E-prime 2 software. The participants sat at a distance of approximately 57 cm from the CRT monitor (90-Hz refresh rate; screen resolution 1024 $\times$ 768) and used a standard mouse to rate the attractiveness of each face. The faces were randomly drawn from a set of 242 faces used in Rhodes, Simmons, and Peters (2005). In total there were 121 Caucasian male faces with a mean age of 23.4 years ($SD = 6.0$, range 18–47 years) and 121 Caucasian female faces with a mean age of 22.9 years ($SD = 5.6$, range 17–51 years). The faces were color photographs taken from the front-view under symmetric and uniform lighting conditions from a distance of 190 cm and were rotated if necessary so that both pupils were aligned horizontally. Face expressions were neutral and a black, oval mask measuring 10 $\times$ 13 degrees of visual angle was placed over each face so that the inner hairline and face outline were visible but most of the hair was covered. The background color was black (< 0.5 cd m^{-2}) and kept constant during the course of the experiment. The attractiveness rating scale consisted of a centrally presented horizontal white (126.4 cd m^{-2}) line that was 201 pixels (6.25°) long and 2 pixels wide with white vertical tick marks (20 pixels in height) at each end. Beneath each tick mark were labels "not attractive" and "attractive" to indicate the boundary conditions of the scale at the left and right ends, respectively. A third white vertical line segment (same dimensions and color as the tick marks) was used as the "attractiveness marker" and it could be moved back and forth along the rating scale using the mouse. A mouse-click recorded the marker's position on the scale to indicate the perceived attractiveness of each face, with the scale ranging from -100 (left-hand end) to $+100$ (right-hand end) pixels for a total length of 201 pixels.

Design and procedure

An illustration of the paradigm used is shown in Figure 1. Each trial started with the presentation of a randomly determined face drawn from the set of 242 face images for a duration of 500 ms. Subsequently, the face was replaced by a screen that was blank except for the attractiveness scale. The position of the attractiveness marker was randomly determined on each trial, and participants adjusted the marker's position on the scale using the mouse to rate the attractiveness of the face and pressed the left mouse button to record it. After the mouse click, a completely blank screen was presented for 500 ms, and then the next trial began with another 500-ms face presentation. There were four experimental blocks of 242 trials each. The random sampling of face order was independent within each block. As we were interested in intertrial effects, the data from the first trial of each block were excluded, leaving a total of $4 \times 241 = 964$ trials per subject. Participants received instructions prior to the experiment and performed 10 practice trials to get familiar with the task of setting the attractiveness marker.

Results and discussion

Distribution of attractiveness ratings as a function of the previous trial's attractiveness

The group mean attractiveness rating for each face in rank order is shown in Figure 2a. It is clear that when the ratings are presented in rank order that they form an orderly collection that varies gradually and smoothly from less than grand mean attractiveness (dashed line) to more than grand mean attractiveness, and does so in a symmetrical way without discontinuities that would indicate a distinct subgroup. The steeper sections at each end indicate some outliers judged as very attractive or very unattractive, as would be expected of any Gaussian-like distribution. Overall, the faces were rated as not being particularly attractive

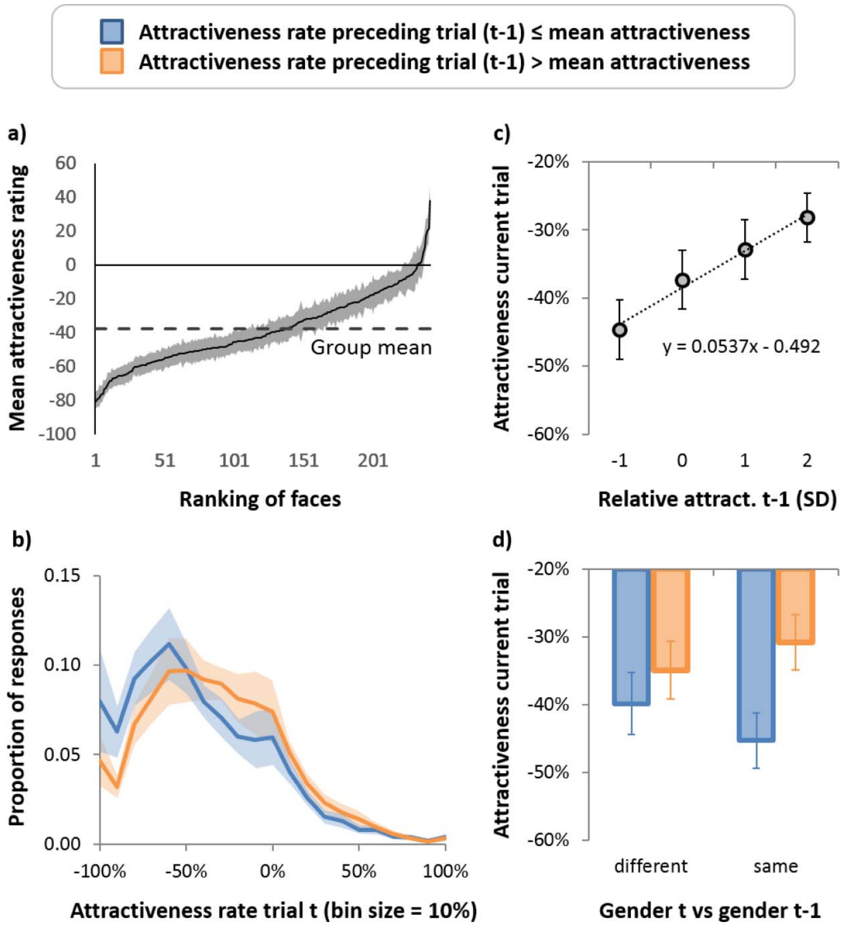

Figure 2. Results of Experiment 1. (a) Group mean rating for each face in rank order. The dashed line indicates the group mean attractiveness rating. (b) Group mean of the attractiveness ratings distributions from 15 observers (bin size = 10%), as a function of whether the face on the preceding trial was rated as being less attractive (blue line) or more attractive (orange line) than the group mean average attractiveness rating computed over all faces. The shaded area represents ±1 standard error of the mean. (c) Attractiveness ratings as a function of the relative attractiveness of the preceding trial's face, with relative attractiveness calculated as deviation from mean attractiveness in units of standard deviation. The dotted line shows the best-fitting linear fit to the relative attractiveness scores. (d) Mean attractiveness ratings as a function of the preceding trial's attractiveness, binned by whether the preceding face's gender was the same as, or different from, the current trial. All error bars show ±1 standard error of the mean.

(mean: −37.7%; similar to Rhodes et al., 2005, who used the same face stimuli). Figure 2a illustrates that the faces are not homogeneous (or else there would be no attractiveness serial dependence). Furthermore, the faces smoothly sample a range of attractiveness around the mean, and there are no clear discontinuities in the curve that might indicate a salient subgroup that drives the effect.

The group mean distribution of attractiveness ratings from Experiment 1 calculated over the 15 observers is shown in Figure 2b, with the data divided into two bins depending on the attractiveness rating in the preceding ($t-1$) trial. The blue distribution shows current-trial (t) attractiveness ratings for which the face presented in the preceding trial ($t-1$) was rated as less attractive than the group mean average over all faces, and the orange distribution shows responses in trial t following faces on trial $t-1$ that were rated as more attractive than average. The group mean of the distribution means gives a mean attractiveness rating over all observers and faces of −37.7%. However, Figure 2b clearly shows the lateral position of the distribution of attractiveness ratings is contingent on the attractiveness of the preceding face. This was statistically confirmed by comparing the means of the blue and orange distributions using a two-tailed repeated-measures t test, which showed a significantly higher mean (−32.8%, orange distribution) when the preceding face was more attractive than average, compared to the mean when the preceding face was less attractive (−42.6%, blue distribution), $t(14) = 6.333$, $p < 0.00005$, $\eta_p^2 = 0.741$.

The tendency to rate a current face as being more attractive when the preceding face was attractive also depended on the preceding face's relative attractiveness. To demonstrate this, the attractiveness data for each participant were binned by how far the preceding face deviated in standard deviation units from mean attractiveness, as plotted in Figure 2c. Reflecting the positive skew in the distributions in Figure 2b (see also Rhodes et al., 2005), the deviations in preceding attractiveness were binned as: −1, 0, 1 and 2 *SD* units. A one-way, repeated-measures analysis of variance (ANOVA) was conducted on mean attractiveness rate with relative attractiveness on the preceding trial as the within-subject variable. The ANOVA yielded a highly significant effect of the preceding face's relative attractiveness, $F(3, 42) = 26.807$, $p < 0.000005$, $\eta_p^2 = 0.657$, indicating that the attractiveness rating on a given trial increases at a rate of 5.3% with increasing deviation of the preceding face from the mean. This was statistically confirmed by separate two-tailed *t* tests (*SD* −1 vs. 0: $t[14] = 7.402$, $p < 0.00001$; *SD* 0 vs. 1: $t[14] = 5.131$, $p < 0.0005$; *SD* 1 vs. 2: $t[14] = 2.177$, $p = 0.047$).

Effect of face gender between current and preceding trials

An interesting question is whether the intertrial attractiveness effect transfers from one gender to another. To investigate this, we conducted a two-way, repeated-measures ANOVA on mean attractiveness ratings with the preceding trial's attractiveness (less vs. more attractive than average) and intertrial gender congruence (same face gender on current and preceding trials vs. different genders) as within-subject variables. The effect of gender on the intertrial attractiveness effect is shown in Figure 2d. The ANOVA yielded a highly significant main effect of the preceding trial's attractiveness, $F(1, 14) = 39.539$, $p < 0.00005$, $\eta_p^2 = 0.739$, as face ratings were higher when the face on the preceding trial was more attractive than average (–32.9%) compared with when it was less attractive (–42.6%). There was no significant main effect of gender congruence between trials, $F(1, 14) = 0.789$, $p = 0.389$, $\eta_p^2 = 0.053$. The interaction between gender congruence and preceding attractiveness was also significant, $F(1, 14) = 14.152$, $p = 0.002$, $\eta_p^2 = 0.503$. Specifically, when the genders on the current and preceding trials matched (Figure 2d, right-hand columns), the positive dependency of current face attractiveness on the preceding face's attractiveness was greater than when face gender did not match between trials (Figure 2d, left-hand columns). This was confirmed by two-tailed repeated-measures *t* tests for each gender condition. When consecutive faces had the same gender, attractiveness ratings were significantly lower when the preceding trial contained a face rated as less attractive than average (–45.3%) compared to when the preceding face was rated more attractive (–30.9%), $t(14) = 7.960$, $p < 0.00001$, $\eta_p^2 = 0.819$. A similar significant intertrial effect was observed when the consecutive faces had the opposite gender, (–39.9% vs. –35.0%), $t(14) = 2.3$, $p = 0.039$, $\eta_p^2 = 0.271$. Another two-tailed *t* test comparing the intertrial effects for same gender faces with different gender faces on two consecutive trials yielded a significant effect, $t(14) = 3.762$, $p = 0.002$, $\eta_p^2 = 0.503$. This confirms that the positive serial dependency for face attractiveness is present regardless of the gender of the faces, and that the significant two-way interaction arises because the dependency is stronger when the gender of consecutive faces is the same (14.4%) compared with different (4.9%).

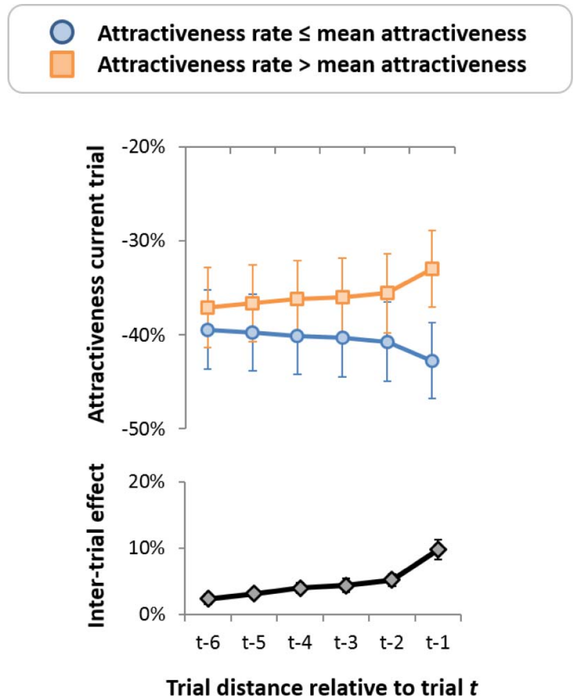

Figure 3. Results of Experiment 1. Upper panel: Mean attractiveness ratings as a function of whether trial $t-1$ back to trial $t-6$ attractiveness ratings were less than (blue symbols) or greater than (orange symbols) the group mean attractiveness calculated over all faces. Lower panel: Intertrial effect (the difference of the upper panel) as a function of the trial distance relative to the current trial *t*. Error bars show ±1 standard error of the mean.

Attractiveness rating as a function of trial distance

An intriguing question is whether the attractiveness rating on a given trial depends solely on the attractiveness of a face on the previous trial, or whether the attractiveness of faces presented several trials back also impacts on the current trial. Figure 3 (upper panel) illustrates the mean attractiveness rating as a function of whether the attractiveness rating was less than (blue circles) or greater than (orange squares) the group mean attractiveness calculated over all faces, and does so for a range of previous trials from trial $t-1$ back to trial $t-6$. The lower panel plots the intertrial effect as a function of trial distance.

We conducted a two-way, repeated-measures ANOVA on mean attractiveness ratings with attractiveness (less vs. more attractive than average) and trial distance ($t-1$ back to $t-6$) as within-subject variables. The ANOVA yielded a significant attractiveness effect, $F(1, 14) = 64.452$, $p < 0.0001$ as well as a significant two-way interaction, $F(5, 70) = 9.633$, $p = 0.0003$. The significant interaction suggests that the intertrial attractiveness effect varies as a function of the trial distance relative to trial t (see Figure 3, lower panel). The interaction was further investigated using two tailed t tests for each trial distance condition. The t test yielded a significant intertrial effect up to five trials back (all p values < 0.0011, Bonferroni corrected), but not for trial $t-6$ ($p = 0.096$, Bonferroni corrected). The results are consistent with recent studies showing that perceived attractiveness on a given trial t not only depends on the face shown on the previous trial, but also on the face two (Taubert, Van der Burg, & Alais, 2016), and three trials back (Taubert, Alais, & Burr, 2016) and even further to five trials back (Xia et al., 2016).

Experiment 1 demonstrates a clear, positive serial dependency for face attractiveness. A face is rated as being more attractive if it is preceded by another face rated highly for attractiveness, and vice versa. This positive or "assimilative" dependency for face attractiveness squares with the findings of Xia et al. (2016) as well as Taubert, Van der Burg, and Alais (2016). Our results also show interesting further insights. Attractiveness ratings for the set of 242 faces used in this experiment were widely distributed (Figure 2b) and we observed that the strength of the attractiveness serial dependency scaled with how deviant the preceding face's attractiveness rating was. That is, the further a preceding face was rated from the mean attractiveness of the face set, the more influence it exerted over the current face. We also found that the attractiveness serial dependency interacted with the gender of the faces, being stronger between consecutive faces of the same gender (whether both male or both female) than for consecutive faces of different gender (Figure 2d). As a change in gender between trials did not eliminate the effect, it could indicate that a response bias underlies the remaining effect, although it is also possible that some aspects of attractiveness that are not gender specific, such as face symmetry and skin condition (Fink, Grammer, & Thornhill, 2001; Rhodes, Proffitt, Grady, & Sumich, 1998) underlie the observed serial dependency.

Experiment 2: Time course of the face attractiveness sequential dependency

The aim of Experiment 2 is to investigate the time course of the intertrial attractiveness effect. Our interest is in finding the minimum exposure duration for the preceding face that will still exert a significant assimilative influence over the attractiveness rating of the current face. A recent study by Xia and colleagues (2016) conducted a similar experiment in which they analyzed the effects of various intertrial intervals between consecutive face stimuli with a fixed 1-s duration. Our experiment adopts the converse approach: We vary the stimulus duration of the previous face, while holding the current face duration and the intertrial interval constant. The design of our experiment involves pairs of adapter and test trials (see also Van der Burg, Alais, & Cass, 2018, for a similar approach), where adapter trials are analyzed as the preceding trial, and test trials are the current trial. Adapter trials have a duration drawn randomly from a set of durations (22, 56, 111, 244, 500 and 1,000 ms) while test trials have a fixed duration of 500 ms, so that an intertrial effect cannot be explained by the current trial's presentation duration. For adapter trials, the short intervals were included to test the boundary condition under which we can observe reliable intertrial effects, whereas a 1-s duration was included to examine whether an asymptote was reached at 500 ms, the duration used in Experiment 1.

Method

Participants

Twenty-three subjects participated in Experiment 2 (14 females; mean age = 22.6 years, range 17–46 years). All participants were paid $AU20 per hour and were naive as to the purpose of the experiment. The number of subjects was increased to provide extra experimental power when the data were binned into different adapter-trial durations.

Experiment 2 was similar to Experiment 1, except for the following changes. In Experiment 2, we divided the trial sequence into test and adapter trials. The adapter and test trials were always presented in alternating order so that a test trial always followed an adapter trial. The presentation duration for the test trials was always 500 ms, whereas the duration for an adapter trial was randomly drawn from a set of stimulus durations: 22, 56, 111, 244, 500, or 1,000 ms. To discount effects of stimulus persistence on adapter trials, each face presentation was followed by a random luminance noise mask for 100 ms. We used a new set of

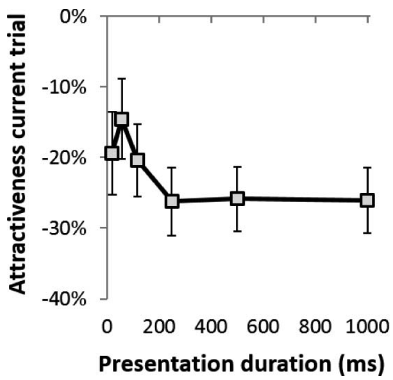

Figure 4. Results from Experiment 2 showing face attractiveness ratings as a function of the presentation duration on adapter trials. Error bars show ±1 standard error of the mean.

36 faces in this experiment, using the same set of 36 faces for test and adapter trials so that any differences in the intertrial effect as a function of duration could not be explained by different sets of faces. Observers completed 72 trials per adapter trial duration, with each of the 36 faces presented twice in independent random orders. There were six durations, making a total of 432 adapter and 432 test trials per observer.

Results and discussion

Attractiveness as a function of presentation duration

Before analyzing the data in terms of the preceding trial, we first plotted the relationship between attractiveness ratings and stimulus duration (Figure 4), as the attractiveness rating on adapter trials may vary as a function of the presentation duration of the faces. For this we used the data obtained from the adapter trials, for which face duration varied among six levels: 22, 56, 111, 244, 500, and 1,000 ms. A one-way repeated-measures ANOVA with face presentation duration as the within-subject variable yielded a highly significant effect of duration, $F(5, 110) = 8.8$, $p < 0.001$, $\eta_p^2 = 0.287$, indicating that the face attractiveness ratings differed across the different presentation durations. As is clear from Figure 4, attractiveness ratings were higher for brief face presentations (22, 56, and 111 ms) and then stabilized at a lower level for longer presentation durations (244, 500, and 1,000 ms), $t(22) = 3.182$, $p = 0.004$. This stabilization was statistically confirmed by a separate ANOVA with face presentation duration (244, 500, and 1,000 ms) yielded no significant effect at all $F(2, 44) = 0.08$, $p = 0.861$, $\eta_p^2 = 0.004$.

Distribution of attractiveness ratings as a function of the previous trial's attractiveness

The group mean distribution of attractiveness ratings for test trials (face presentation duration is 500 ms) is shown in Figure 5a, with the data divided into two bins depending on the attractiveness rating in the preceding trial. As in Experiment 1, Figure 5a clearly shows the lateral position of the distribution of attractiveness ratings is contingent upon the attractiveness of the face on the preceding adapter trial. This was statistically confirmed by a two-tailed t test, $t(22) = 6.527$, $p < 0.00001$, $\eta_p^2 = 0.659$, indicating a significantly higher mean (–21.5%, orange distribution) when the preceding face was more attractive than average, compared to the mean when the preceding face was less attractive (–29.5%, blue distribution).

The time course of the inter-trial attractiveness effect

Figure 5b (upper panel) illustrates how attractiveness ratings for test trials (for which the presentation duration was always 500 ms) varied as a function of the preceding face's attractiveness and its presentation duration. Here, a face on the preceding trial was considered as being less or more attractive when the participants rated the face lower or higher than the overall mean attractiveness. The mean attractiveness was calculated for each face presentation duration separately, provided that the face attractiveness ratings vary with presentation duration (see Figure 4). For instance, if the preceding trial duration was 56 ms, we used the average attractiveness rating for 56-ms presentations when binning that trial's attractiveness rating into "less than average" or "more than average." The lower part of Figure 5b (black line) shows the difference between the orange and blue plots above, and thus shows the magnitude of the intertrial attractiveness effect as a function of presentation duration. This intertrial effect was examined using a one-way repeated-measures ANOVA with the preceding trial's presentation duration as the within-subject variable. This yielded a significant linear contrast, $F(1, 22) = 8.253$, $p = 0.009$, $\eta_p^2 = 0.273$, indicating that the intertrial effect increased with increasing presentation duration. No other higher order contrasts were significant (all Fs < 2.772, $p > 0.110$). Separate two-tailed t tests comparing the magnitude of the intertrial effect against zero yielded significant intertrial effects for all durations of the preceding trial that were 56 ms or greater, all $t(22) > 2.9$, all p values < 0.007. The intertrial effect was not significant when the presentation duration was 22 ms, $t(22) = 1.2$, $p = 0.255$.

The results from Experiment 2 replicate the strong assimilative serial dependency for face attractiveness seen in Experiment 1 by which a face is rated more attractive if preceded by another face rated highly for attractiveness, and vice versa. Before discussing the intertrial analyses, the first finding of note was that face attractiveness ratings depended on exposure duration (Figure 4). For very brief face stimuli, attractiveness

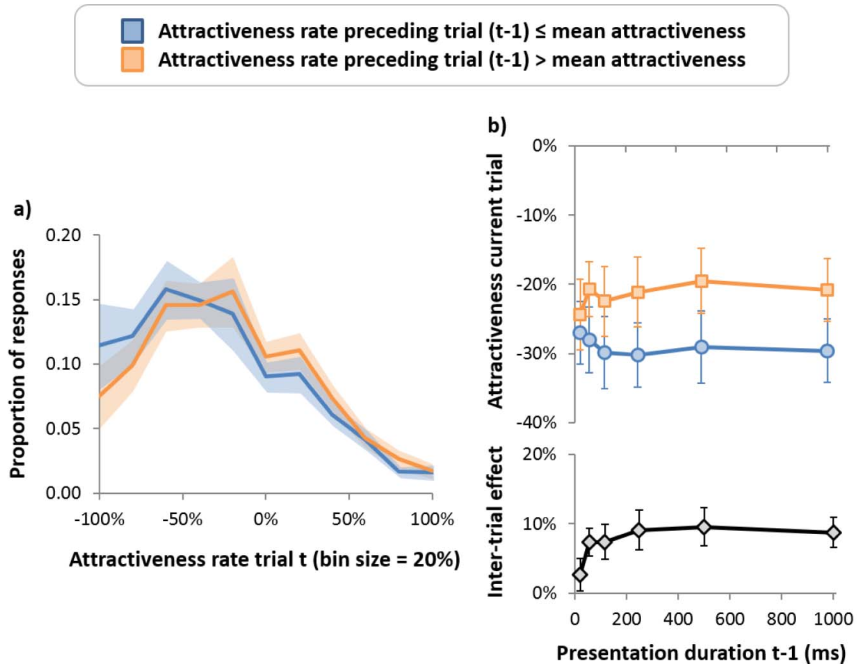

Figure 5. Results of Experiment 2. (a) Group mean of the attractiveness ratings distributions for test trials (bin size = 20%), as a function of whether the face on the preceding adapter trial was rated as being less attractive (blue line) or more attractive (orange line) than the group mean average attractiveness rating computed over all faces. The shaded area represents ±1 standard error of the mean. (b) Upper panel: Mean attractiveness ratings as a function of the previous adapter trial's presentation duration, as a function of whether the previous trial's attractiveness rating was less than (blue symbols) or greater than (orange symbols) the group mean attractiveness calculated over all faces with the same durations. Lower panel: Intertrial effect (the difference of the upper panel) as a function of presentation duration on the preceding adapter trial. Error bars show ±1 standard error of the mean.

ratings were higher than those presented for longer durations. This bias toward higher attraction for briefly seen faces appears to be a very transient effect, as attractiveness ratings stabilized for presentations of 244 ms and greater, but were clearly elevated above the stabilized values the three briefest durations tested here: 22, 56, and 111 ms. To our knowledge, this is the first time a relationship between face attractiveness and stimulus brevity has been reported and this effect is one that merits further investigation.

The specific aim of Experiment 2 was to examine the time course of the attractiveness serial dependency by manipulating the exposure duration of adapter trial faces and quantifying their assimilative influence over test trial faces of a standard duration, with intertrial interval left constant. The results are shown in Figure 5, with the magnitude of the serial dependency shown in the lower panel of Figure 5b. This plot shows that the previous stimulus can be very brief and still exert a significant assimilative pull over the subsequent stimulus, with all adapter trial durations producing significant shifts in test trial attractiveness, except for the shortest adapter duration of 22 ms. Although we did not test longer durations, it is likely this would lead to a repulsive serial effect. This is expected based on numerous adaptation studies showing typical repulsive perceptual aftereffects occur following several seconds of face adaptation (Burton, Jeffery, Bonner, & Rhodes, 2016; Leopold et al., 2005), although, consistent with our findings here, subsecond face adaptation has been reported to produce positive effects (Oruç & Barton, 2010). This point is discussed further in the General discussion.

Finally, our data make an interesting comparison with the findings of Xia and colleagues (2016). Their experiment was an online study of face attractiveness that controlled stimulus presentation times at 1 s but response times were self-paced and varied widely. When responses were binned into 1-s intervals, they found the positive serial dependency for attractiveness was significant for up to 6 s after stimulus presentation with no clear tendency to decay over this period. This shows that once induced by a preceding stimulus, the assimilative pull over the current stimulus can endure for many times longer than the priming stimulus' exposure duration. Our approach to exploring the time-course of the effect was to vary the duration of the priming stimulus. This revealed that the serial assimilative effect is induced very quickly, with 56 ms being sufficient to produce a serial dependency equivalent in magnitude to that induced by a 1,000-ms exposure. Both these observations are consistent with the serial dependency effect for face attractiveness not depending on sensory adaptation, a point taken up in the General discussion.

Experiment 3: Is the sequential dependency for face attractiveness task dependent?

In Experiment 3 we test whether the sequential dependencies, as observed in Experiments 1 and 2, are perceptual in nature or require an explicit attractiveness judgment. A recent study by Ritchie et al. (2017) reported that when participants were learning new identities for the purpose of future recognition, they automatically formed impressions of each person's attractiveness, and that once formed, these impressions could not be suppressed. Even though participants were explicitly instructed to rate each specific image independent of previously learned images of that person, they could not inhibit their prior impressions of each person's attractiveness. If the observed intertrial attractiveness effects we report are due to a perceptual effect, then we expect that simply viewing a face should automatically induce an intertrial effect as participants would automatically generate an impression of attractiveness. In other words, manipulating the task should not eliminate the intertrial effect. To investigate whether the intertrial effect requires an explicit attractiveness judgment, we conducted an experiment consisting of two parts. The first part was a single-task condition in which participants made an explicit attractiveness judgement on every trial so that we could replicate Experiment 1's intertrial attractiveness effect and obtain estimates for each face's attractiveness rating. The second part used the same set of faces but followed an alternating, two-task format. On adapter trials, participants judged the gender of the face, and on test trials, they made attractiveness ratings. In the second part of the experiment, if viewing a face alone is enough to trigger an intertrial attractiveness effect, we expected to observe a similar sequential effect to that observed in the single-task condition. If, however, it is important to explicitly make an attractiveness judgment on the preceding trial in order to generate an attractiveness assimilation effect, the alternating two-task condition should not produce any attractiveness assimilation effect on the current trial as the previous task will always be a gender task.

Method

Participants

Fourteen students (10 females; mean age = 22.5 years, range 18–33 years) participated in the experiment, based on our previous research (Taubert, Van der Burg, & Alais, 2016). Participants were paid $AU20 per hour and were naive as to the purpose of the experiment.

The experiment was identical to Experiment 1, except for the following changes. The experiment consisted of two different parts. In the first part, participants saw 242 random faces and performed a single task, which was to rate the attractiveness of each face (as in Experiments 1 and 2). The first part was identical to Experiment 1, except that there was only one experimental block of 242 trials, instead of four as in Experiment 1. In the second part of the experiment, participants saw the same 242 faces but performed two different tasks in alternating order. On adapter trials, participants judged the gender of the face by pressing the 'm' or 'f' key to indicate whether they thought the face was male or female, respectively. On test trials, participants performed the attractiveness rating task, as in Experiments 1 and 2. Participants first performed the single task condition (Part 1), and subsequently completed the alternating dual-task condition (Part 2). All faces were presented for 500 ms.

Results and discussion

Distribution of attractiveness ratings as a function of the preceding trial's attractiveness

The group mean distribution of attractiveness ratings from the single-task condition is shown in Figure 6a, with the data divided into two bins depending on the attractiveness rating in the preceding trial. The attractiveness ratings from the alternating two-task condition are plotted in Figure 6b according to the same format but with one difference: The data were divided into two bins depending on the *estimated* attractiveness of the preceding trial's face. This was necessary because the preceding trial in the alternating-task condition was always the gender task. The attractiveness for these trials was estimated for each face and observer using the attractiveness ratings provided in the single-task condition (where exactly the same faces were shown).

The group mean of the distribution means gives a mean attractiveness rating over all observers and faces of −33.0% in the single-task condition and −31.2% in the alternating-task condition. A two-way, repeated-measures ANOVA was conducted on mean attractiveness rate with task (single vs. alternating) and the preceding trial's attractiveness (less than or greater than average) as within-subject variables, with the results plotted in Figure 6c. The ANOVA yielded no significant difference between the single- and alternating-task conditions, $F(1, 13) = 0.199$, $p = 0.663$, $\eta_p^2 = 0.015$. There was a significant interaction between task and preceding attractiveness, $F(1, 13) = 14.911$, $p =$

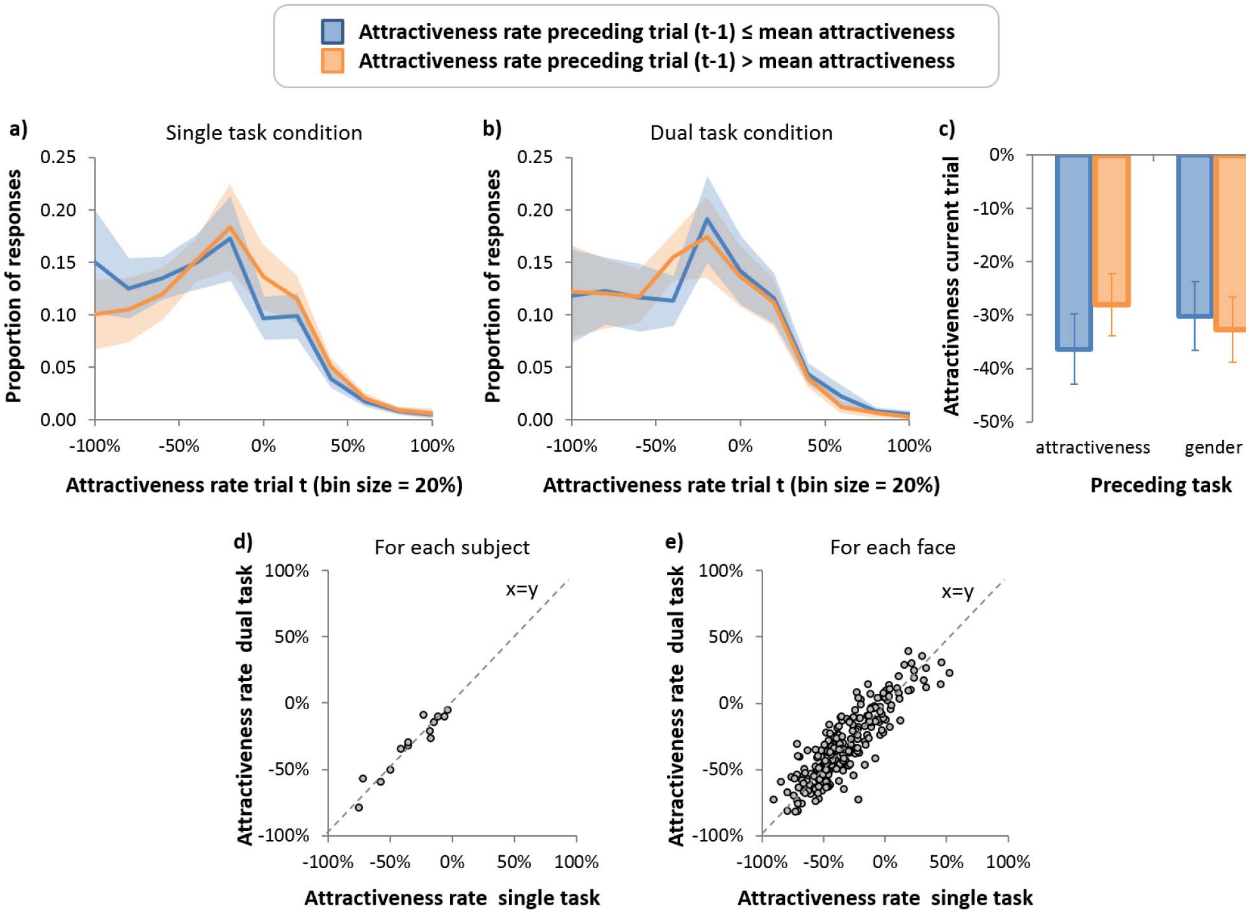

Figure 6. Results of Experiment 3. (a) Group mean of the attractiveness ratings distributions in the single-task condition, binned by whether the face on the preceding trial was rated as less attractive (blue line) or more attractive (orange line) than average. (b) Group mean of the attractiveness ratings distributions in the alternating-task condition, binned by whether the face on the preceding trial was rated as being less attractive (blue line) or more attractive (orange line) than the estimated average. (c) Mean attractiveness ratings as a function of the preceding trial's attractiveness and the task to perform. (d) Scatterplot showing the correlation for all 14 observers between their mean attractiveness ratings for the entire face set in the single-task and the alternating-task conditions. (e) Scatterplot showing the correlation between the mean attractiveness rating given to each individual face in the single-task and alternating-task conditions (collapsed over all participants).

0.002, $\eta_p^2 = 0.534$, indicating that the intertrial attractiveness effect depended on the task in the preceding trial. The interaction was further examined for each task condition by two separate two-tailed t tests, which showed a highly significant intertrial attractiveness effect in the single-task condition, $t(13) = 5.6$, $p < 0.0001$, $\eta_p^2 = 0.711$, but not in the alternating-task condition, $t(13) = 1.1$, $p < 0.298$, $\eta_p^2 = 0.083$. In other words, the serial dependency for face attractiveness was observed when participants performed an attractiveness judgment on the preceding trial, but not when they performed a gender judgment task on the preceding trial.

A potential concern might be that our conclusion is weakened in being based on a comparison between the actual attractiveness ratings in the single-task condition and the attractiveness ratings in the alternating-task condition (where adapter trials involved a judgment of face gender rather than an attractiveness rating). However, we observed a very strong correlation between attractiveness ratings made in the single- and alternating-task conditions across all participants. The scatterplot in Figure 6d shows this relationship for each observer's mean attractiveness ratings over the entire set of faces and produced a Pearson correlation of 0.956, $p < 0.000001$, indicating very consistent attractiveness ratings within observers between the single- and alternating-task conditions. We also analyzed the consistency of face attractiveness ratings between single- and alternating-tasks for each face in the set (averaged over participants) and again found a very strong relationship (Pearson correlation $= 0.869$, $p < 0.000001$). Both of these strong correlations support the argument that using face attractiveness ratings obtained in the single-task condition to estimate attractiveness in the gender-judgment trials of the alternating-task condition was a valid procedure and cannot explain the lack of an assimilative effect in the alternating-task condition. Therefore, we conclude that the intertrial attractiveness effect is task dependent and requires an explicit attractiveness judgment to be performed in the preceding trial in order to influence the current trial.

The aim of Experiment 3 was to test whether the positive serial dependency for face attractiveness that was well established in Experiments 1 and 2 is a perceptual effect or requires an explicit attractiveness judgment. Comparing the results of two conditions where test and adapter trials either both required an attractiveness rating (the single-task condition) or required an alternating gender/attractiveness task over test and adapter trials, the results show clearly that the intertrial attractiveness effect does depend on the task in the preceding trial. If simply viewing a face on a given trial were sufficient to produce an assimilative influence over the subsequent trial, it should have occurred in the alternating-task condition as well, as all trials involved face stimuli (and, indeed, the faces were the same in both conditions). The fact that no assimilative effect was observed in the alternating-task condition—when the only difference was the task required on the adapter trials—may suggest that the intertrial attractiveness effect is task dependent and presumably not due to a perceptual process.

General discussion

In a series of experiments, we demonstrate a positive serial dependency for face attractiveness. Faces were consistently rated as more attractive if preceded by another face rated highly for attractiveness (and vice versa), consistent with the assimilative dependency reported by Xia et al. (2016) for face attractiveness. The strength of the assimilative dependency increased as the preceding face deviated further from mean attractiveness and was stronger for consecutive faces of the same gender (Experiment 1). The minimum exposure duration (Experiment 2) required in the previous face to produce the assimilative dependency was surprisingly brief, with a 56-ms exposure sufficient to produce a significant intertrial effect. Finally, we found evidence that the serial dependency for face attractiveness is not automatic but is task dependent (Experiment 3). Simply viewing a face on a given trial was not sufficient to produce an assimilative influence: The face had to be rated for attractiveness in order to elicit an attractive influence over the subsequent trial.

The finding of a positive serial dependency for face attractiveness fits with the proposed "continuity field" framework proposed by Fischer and Whitney (2014). The notion is that much of the external world that provides the visual system's input is stable and thus momentary fluctuations can be discounted by averaging over time to boost signal-to-noise ratio with little cost to perceptual veracity (Kiyonaga et al., 2017). The attractiveness of a face is a salient and stable characteristic that rarely changes spontaneously so that fluctuations from moment to moment are more likely due to changes in other variables, such as viewpoint, lighting, distance, etc. While this would apply to faces of the same or similar identity, the faces in our study were never repeated from trial to trial, and so rather than being an identity-specific effect, we argue that our findings indicate a more general effect, such as priming (Oruç & Barton, 2010) or short-term averaging, which the brain uses to reduce noise and improve perceptual precision (Cicchini et al., 2018). Face attractiveness is therefore an attribute that could benefit from integration over the recent past (over the last five trials, for instance) and thus manifest as a positive serial dependency, as reported by Xia et al. (2016) and in related studies that reported assimilative biases for sequential face attractiveness judgments using Likert scales (Kondo, Takahashi, & Watanabe, 2013; Kramer, Jones, & Sharma, 2013).

An outstanding question arising from those earlier observations was whether the assimilative effect was driven postperceptually by the previous response, similar to an anchoring effect (DeCarlo & Cross, 1990; Tversky & Kahnemen, 1974; Ward & Lockhead, 1971), or by the previous stimulus (Liberman et al., 2014). The possibility of serial dependencies being due to response bias was addressed specifically by Xia et al. (2016) for face attractiveness. They had participants do two blocks of their face attractiveness rating procedure, using the same set of faces but in independent orders in each block. This allowed them to test for serial dependency in one run using the attractiveness ratings from the other, independent run. Xia et al. argued that the current face's attractiveness rating could not be driven by the response to the previous face as the attractiveness rating for the previous face came from an independent rating in a second run that used the same faces but in a randomly shuffled order. By decorrelating attractiveness ratings and responses in this way, Xia et al. concluded that response bias could not explain their positive serial dependency effect. Other studies have addressed this question and also concluded there is a strong perceptual component in serial dependency (Cicchini, Mikellidou, & Burr, 2017). In our study, several observations argue against a response-bias account. In Experiment 2 there was no serial dependency for very short stimulus presentations, even though the task and rating response remained unchanged. Brief durations (e.g., our 22-ms condition) should reveal a greater influence of response bias as stimulus salience and adaptation would be expected to be less with so little enough time to create an impression of face attractiveness (Pegors et al., 2015). In Experiment 1, we found stronger serial dependence when the face gender was the same (compared to different) on the current and previous trials, which cannot be explained by response bias as the task and response keys remained the same.

It is important to note that the attractiveness dependency was reduced in Experiment 1 by faces of different gender across successive trials, but was not eliminated. This leaves the possibility that at least a part of the intertrial effect may be due to decision-level factors, or alternatively that the residual attractiveness dependence for different gender faces is in fact due to facial features that are not gender specific, such as symmetry and skin condition (Fink et al., 2001; Rhodes et al., 1998). Experiment 3 could be interpreted as supporting the response bias account, as we found no serial dependency for face attractiveness when participants performed a different task (gender discrimination) on the preceding trial. If attractiveness is an attribute that is encoded automatically upon seeing a face, as recent evidence suggests (Ritchie et al., 2017), then the positive dependency would have been maintained despite the change of task. Some caution is needed before concluding that this endorses a response-bias account. In the Ritchie et al. study, participants were shown a face for 5 s in order to memorize it and therefore had enough time to process more attributes than identity alone. Moreover, attractiveness may have been encoded strategically as another dimension on which to individuate individual faces. Further research would be needed to assess whether attractiveness is "automatically" encoded in brief presentations, such as used here, and thus to evaluate the task dependency seen in our dual-task condition. Overall, however, it seems reasonable that there would be a component due to response. This is consistent with recent work showing that while serial dependence effects do occur at the time of perception and do contain a strong stimulus-driven component (Cicchini et al., 2017; Manassi et al., 2018), they also have a component due to response (Cicchini et al., 2017), which tends to add to the stimulus effect. Together, the evidence suggests a role for decision-level influences in serial dependence but that it cannot entirely account for the findings. Our results from Experiments 1 and 2 regarding exposure duration and gender change over successive trials are consistent with this interpretation.

Another relevant factor in explaining the intertrial attractiveness effects is the dependency on exposure duration. Experiment 2 showed that the assimilative dependency was evident after surprisingly brief stimulus exposures, with presentations of 56 ms sufficient to produce a significant effect. This suggests the effect is not likely to be due to sensory adaptation, for two reasons. First, previous research has shown that the size of aftereffects increases with the duration of stimulus exposure (Kohn, 2007; Webster, 2003) and second, perceptual aftereffects resulting from adaptation to various face attributes have generally been reported to be repulsive (i.e., a negative dependency: Leopold et al., 2005). Studies examining the time course of adaptation have observed repulsive effects for identity and expression aftereffects for adaptation durations down to 1 s (Burton et al., 2016; Leopold et al., 2005). Although this appears clear for exposure durations of 1 s or more, one study that used very short (subsecond) adaptation times similar to ours (Oruç & Barton, 2010) reported attractive aftereffects for face recognition. Our data (Figure 5) agree with this, clearly showing an attractive (positive) serial dependency, rather than a repulsive effect, for face attractiveness and show no growth in effect size over a four-fold increase in presentation duration from 244 to 1,000 ms. As well as the lack of growth over longer durations, there is a strikingly rapid onset of the effect with only the 22-ms exposure failing to produce a significant effect. Rather than early sensory activity, this is consistent with a more central process involving memory where once the visual object is encoded and recognized it can exert an attractive influence on subsequent perception (Bliss et al., 2017). On this view, presentation times only need to be long enough to ensure encoding in memory and beyond that longer durations are equivalently effective as an assimilative prime for the following stimulus. The task dependency observed in Experiment 3 squares with this view as only the task-relevant information would be retained in working memory on a given trial (e.g., gender) and so would not influence the orthogonal task on the subsequent trial (e.g., attractiveness). This priming role would hold for a relatively short period before increasing sensory adaptation counteracts it and eventually produces an effect of opposite sign (Fischer & Whitney, 2014; Kiyonaga et al., 2017). Consistent with this, Fischer and Whitney found that the serial dependency for orientation obtained with 500-ms presentations was positive but was negative for 5000-ms presentations—a repulsive effect typical of aftereffects from sustained adaptation.

To recap, we observed that face attractiveness exhibits a positive serial dependency. More specifically, we found that stimuli as brief as 56 ms produce a significant positive dependency similar in magnitude to that produced by stimuli presented for 1,000 ms. We observed stronger positive dependencies between same-gender faces, and found a task dependency: Alternating gender discrimination trials with attractiveness rating trials produced no serial dependency. We conclude that a perception-stabilizing assimilation effect operates in face attractiveness perception that is task dependent and is acquired surprisingly quickly. In line with other recent findings (Cicchini et al., 2017; Xia et al., 2016), we argue this effect has a strong perceptual component, although we cannot discount postperceptual processes entirely.

Keywords: face perception, aftereffects, vision

Acknowledgments

Supported by a grant from the Australian Research Council (DP150101731) to DA.

Commercial relationships: none.
Corresponding author: Erik Van der Burg.
Email: vanderburg.erik@gmail.com.
Address: Department Experimental and Applied Psychology, Vrije Universiteit Amsterdam, The Netherlands.

References

Alais, D., & Burr, D. (2004). The ventriloquism effect results from near-optimal bimodal integration. *Current Biology*, *14*(3), 257–262.

Alais, D., Kong, G., Palmer, C., & Clifford, C. (2018). Eye gaze direction shows a positive serial dependency. *Journal of Vision*, *18*(4):11, 1–12, https://doi.org/10.1167/18.4.11. [PubMed] [Article]

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.

Alais, D., Orchard-Mills, E., & Van der Burg, E. (2015). Auditory frequency perception adapts rapidly to the immediate past. *Attention, Perception & Psychophysics*, *4*(1), 896–906.

Alexi, J., Cleary, D., Dommisse, K., Palermo, R., Kloth, N., Burr, D., & Bell, J. (2018). Past visual experiences weigh in on body size estimation. *Scientific Reports*, *8*(1), 215.

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.

Burton, N., Jeffery, L., Bonner, J., & Rhodes, G. (2016). The timecourse of expression aftereffects. *Journal of Vision*, *16*(15):1, 1–12, https://doi.org/10.1167/16.15.1. [PubMed] [Article]

Cicchini, G. M., Anobile, G., & Burr, D. (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.

Cicchini, G. M., & Burr, D. C. (2018). Serial effects are optimal. *Behavioural Brain Sciences*, *41*:e229.

Cicchini, G. M., Mikellidou, K., & Burr, D. (2017). Serial dependencies act directly on perception. *Journal of Vision*, *17*(14):6, 1–9, https://doi.org/10.1167/17.14.6. [PubMed] [Article]

Cicchini, G. M., Mikellidou, K., & Burr, D. (2018). The functional role of serial dependence. *Proceedings of the Royal Society B*, *285*(1890):20181722.

Corbett, J. E., Fischer, J., & Whitney, D. (2011). Facilitating stable representations: Serial dependence in vision. *PLoS One*, *6*(1):e16701.

DeCarlo, L. T., & Cross, D. V. (1990). Sequential effects in magnitude scaling: Models and theory. *Journal of Experimental Pscyhology: General*, *119*(4), 375–396.

Dong, D. W., & Atick, J. (1995). Statistics of natural time-varying images. *Network: Computation in Neural Systems*, *6*(3), 345–358.

Ernst, M. O., & Banks, M. S. (2002, January 24). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*, 429–433.

Fink, B., Grammer, K., & Thornhill, R. (2001). Human (Homo sapiens) facial attractiveness in relation to skin texture and color. *Journal of Comparative Psychology*, *115*(1), 92–99.

Fischer, J., & Whitney, D. (2014). Serial dependence in visual perception. *Nature Neuroscience*, *17*, 738–743.

Fox, C. J., & Barton, J. J. S. (2007). What is adapted in face adaptation? The neural representations of expression in the human visual system. *Brain Research*, *1127*, 80–89.

Fritsche, M., Mostert, P., & de Lange, F. P. (2017). Opposite effects of recent history on perception and decision. *Current Biology*, *27*(4), 590–595.

Hsu, S., & Young, A. (2004). Adaptation effects in facial expression recognition. *Visual Cognition*, *11*(7), 871–899.

Keane, B., Bland, N. S., Matthews, N., Carroll, T. J., & Wallis, G. (in press). Rapid recalibration of temporal order judgements: Response bias accounts for contradictory results. *European Journal of Neuroscience*, https://doi.org/10.1111/ejn.14551.

Kiyonaga, A., Scimeca, J. M., & Whitney, D. (2017). Serial dependence across perception, attention, and memory. *Trends in Cognitive Sciences*, *21*(7), 493–496.

Kohn, A. (2007). Visual adaptation: Physiology, mechanisms, and functional benefits. *Journal of Neurophysiology*, *97*, 3155–3164.

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*:160685, https://doi.org/10.1098/rsos.160685.

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

Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level after-effects. *Nature Neuroscience*, *4*, 89–94.

Leopold, D. A., Rhodes, G., Muller, K. M., & Jefferey, L. (2005). The dynamics of visual adaptation to faces. *Proceedings of the Royal Society B*, *272*(1566), 897–904.

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

Manassi, M., Liberman, A., Kosovicheva, A., Zhang, K., & Whitney, D. (2018). Serial dependence in position occurs at the time of perception. *Psychonomic Bulletin & Review*, *25*(6), 2245–2253.

Olson, I. R., & Marshuetz, C. (2005). Facial attractiveness is appraised in a glance. *Emotion*, *5*(4), 498–502.

Oruç, I., & Barton, J. J. S. (2010). A novel face aftereffect based on recognition contrast thresholds *Vision Research*, *50*, 1845–1854.

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 Pscyhology: General*, *144*(3), 664–673.

Rhodes, G., & Jeffery, L. (2006). Adaptive norm-based coding of facial identity. *Vision Research*, *46*, 2977–2987.

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

Rhodes, G., Proffitt, F., Grady, J. M., & Sumich, A. (1998). Facial symmetry and the perception of beauty. *Psychonomic Bulletin & Review*, *5*, 659–669.

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

Ritchie, K. L., Palermo, R., & Rhodes, G. (2017). Forming impressions of facial attractiveness is mandatory. *Scientific Reports*, *7*, 469.

Roseboom, W. (2019). Serial dependence in timing perception. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(1), 100–110.

Suárez-Pinilla, M., Seth, A. K., & Roseboom, W. (2018). Serial dependence in the perception of visual variance. *Journal of Vision*, *18*(7):4, 1–24, https://doi.org/10.1167/18.7.4. [PubMed] [Article]

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 Belle, G., Vanduffel, W., Rossion, B., & Vogels, R. (2015). The effect of face inversion for neurons inside and outside fMRI-defined face-selective cortical regions. *Journal of Neurophysiology*, *113*(5), 1644–1655.

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, https://doi.org/10.1038/srep22740.

Tversky, A., & Kahnemen, D. (1974, September 27). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131.

Van der Burg, E., Alais, D., & Cass, J. (2013). Rapid recalibration to asynchronous audiovisual stimuli. *Journal of Neuroscience*, *33*(37), 14633–14637.

Van der Burg, E., Alais, D., & Cass, J. (2015). Audiovisual temporal recalibration occurs independently at two different time scales. *Scientific Reports*, *5*, 14526, https://doi.org/10.1038/srep14526.

Van der Burg, E., Alais, D., & Cass, J. (2018). Rapid recalibration to audiovisual asynchrony follows the physical—not the perceived—temporal order. *Attention, Perception & Psychophysics*, *80*(8), 2060–2068.

Van der Burg, E., & Goodbourn, P. T. (2015). Rapid, generalized adaptation to asynchronous audiovisual speech. *Proceedings of the Royal Society B*, *282*(1804):e20143083.

Ward, L., & Lockhead, G. (1971). Response system processes in absolute judgment. *Perception & Psychophysics*, *9*(18).

Webster, M. A. (2003). Pattern selective adaptation in color and form perception. *Visual Neuroscience*, *2*, 936–947.

Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004, April 1). Adaptation to natural facial categories. *Nature*, *428*, 557–561.

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science*, *17*(7), 592–598.

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

Yovel, G., & Kanwisher, N. (2005). The neural basis of the behavioral face-inversion effect. *Current Biology*, *15*(24), 2256–2262.