

# Influence of gender membership on sequential decisions of face attractiveness

Aki Kondo · Kohske Takahashi · Katsumi Watanabe

Published online: 14 September 2013  
© Psychonomic Society, Inc. 2013

**Abstract** Responses in a current trial are biased by the stimulus and response in the preceding trial. In a mixed-category sequence, the sequential dependency is weaker when the stimuli of the current and preceding trials fall under different categories. In the present study, we investigated the influence of the gender membership of faces on the sequential dependency. Forty-eight pictures of male and female faces were presented successively. Participants rated the attractiveness, roundness, or intelligence of each face on a 7-point scale. The sequential effect was robustly observed, irrespective of the property to be judged. However, between-gender sequential dependency was weaker than within-gender dependency only in the attractiveness judgment. These findings suggest that the gender of faces serves as a cue for forming category representations when face attractiveness is of interest, and hence that the formation of categories in sequential decisions is an adaptive process that depends on the property to be judged.

**Keywords** Sequential effects · Facial attractiveness · Gender

When we make decisions one by one for sequentially presented items, each decision is biased by the preceding decisions and the preceding items. This phenomenon is known as the *sequential effect* (Holland & Lockhead, 1968). For example, when we estimate the brightness of visual flashes, the ratings on the present stimulus are contrasted to the brightness of preceding stimuli and are also assimilated toward the ratings made on the preceding stimuli. The sequential effect has been widely and robustly

observed with the physical properties of sensory stimuli, such as the loudness and brightness of sensory stimuli or the size of visual objects (Holland & Lockhead, 1968; Petzold & Haubensak, 2001; Ward & Lockhead, 1970). Also, the sequential effect has been found with more complex, nonphysical properties, such as price, performance in sports, auditions for the *American Idol* television show, and cooperativeness in a series of strategic choice decisions (i.e., prisoner's dilemma games; Damisch, Mussweiler, & Plessner, 2006; Matthews & Stewart, 2009; Page & Page, 2010; Vlaev & Chater, 2007). Recently, it has been shown that the sequential effect is observed even for subjective decisions on the properties of items whose external true value is undefined; Kondo, Takahashi, and Watanabe (2012) investigated the sequential effect in face attractiveness judgments and found that the attractiveness ratings were biased toward those on the preceding trials (i.e., response–response sequential effect).

In the real life, it is frequent that we sequentially evaluate various values among various categories. Several studies have suggested that the between-category sequential dependency is weaker than the within-category sequential dependency (Petzold & Haubensak, 2004; Ward, 1985). Ward (1985) presented lights and sounds sequentially, and the participants estimated the brightness of the light or the loudness of the sound. The biasing effects of the preceding stimuli were weaker when the preceding and current stimuli were presented in different modalities (i.e., light–sound or sound–light) than when the stimuli were from the same modality (i.e., light–light or sound–sound). Petzold and Haubensak (2004) investigated the effects of categorical modulations of visual stimuli on the sequential effect. They presented filled or empty squares sequentially, and the participants estimated the sizes of the squares. One experimental group was explicitly instructed to treat the filled and empty squares separately, whereas the other group was instructed to ignore the filling style. The sequential dependency decreased when the filled and empty squares

---

A. Kondo · K. Takahashi · K. Watanabe (✉)  
Research Center for Advanced Science and Technology,  
The University of Tokyo, 4-6-1, Komaba, Meguro-ku,  
Tokyo 1538904, Japan  
e-mail: kw@fennel.rcast.u-tokyo.ac.jp

were seen as different categories. Thus, the category membership of successive items could be one determinant of the magnitude of the sequential effect.

In the present study, we aimed to examine how category representation in sequential judgments is formed. For this, we measured the influence of gender membership on sequential dependency in various types of judgment of faces. Faces can be judged on multiple dimensions; we can also make both subjective and objective evaluations for a face, such as age, gentleness, attractiveness, intelligence, skin color, roundness, and so forth. Does categorical (e.g., gender) membership affect the sequential dependency of various types of judgments with the same magnitude? There might be no obvious reason to take into account the face's gender when an observer is evaluating the roundness of the face. However, it would be possible that the face's gender is automatically considered when the observer is evaluating the attractiveness of the face. In three experiments, we examined the between- and within-gender sequential dependencies in face attractiveness, face roundness, and face intelligence judgments.

### Experiment 1: Face attractiveness judgment

In Experiment 1, we examined whether the magnitude of the sequential effect in face attractiveness judgments would depend on the gender memberships of the current and preceding faces, by reanalyzing the data acquired in Experiment 1 of an earlier study (Kondo et al., 2012).

#### Method

A group of 29 undergraduate students (10 males and 19 females) participated in the experiment as paid volunteers. The stimuli were grayscale pictures of Japanese male and female faces ( $3.3^\circ$  width  $\times$   $4.2^\circ$  height). The individual heads were photographed under conditions that controlled for distance, angle, and luminosity. They all had a neutral expression, varied hairstyles, and were free of glasses. The faces, necks, and hair were all visible in the pictures. For each trial, one of the 48 pictures (24 females and 24 males) was presented. Participants were asked to rate the attractiveness of each face by clicking a mouse on a 7-point scale shown below the face stimulus, and the next stimulus appeared immediately after the participant's response. Each participant completed four blocks consisting of 24 trials each. All pictures were presented twice in the sequence, and the order of stimulus presentation was randomized.

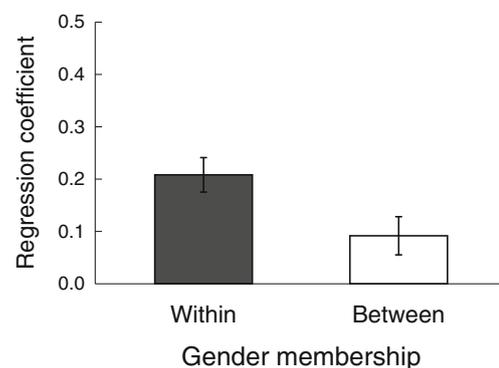
In data analysis, we first divided all trials into two groups on the basis of gender membership between the current and 1-back trials (except for the first trials). One group consisted of within-gender trials, on which the gender of the current face stimulus was the same as that on the immediately preceding trial. The

other consisted of between-gender trials, on which the genders of the faces on the current and immediately preceding trials were different. Each group was submitted to a regression analysis separately. In the regression analysis, the dependent variable was the rating score on the current trials, whereas the independent variable was that on the 1-back trials. The regression analysis was performed for each participant, and then the regression coefficients were averaged among the participants. The positive coefficients would indicate that the rating scores of current trials were assimilated to those of the 1-back trials.

#### Results and discussion

The mean rating scores of attractiveness judgments were significantly higher for the female than for the male faces (male, mean = 3.27,  $SD$  = 0.65; female, mean = 3.98,  $SD$  = 0.63) [ $t(28) = -5.91$ ,  $p < .0001$ ]. Since the difference of mean rating scores between genders would lead to overestimation and underestimation biases in the within-gender and between-gender sequential effects, respectively, we normalized the rating scores of male and female faces separately prior to the regression analysis. Figure 1 shows the averaged regression coefficients of the within- and between-gender groups. We found significant sequential effects in both the within- and between-gender groups. Specifically, the regression coefficients were positive in both groups [within,  $t(28) = 6.32$ ,  $p < .0001$ ; between,  $t(28) = 2.52$ ,  $p < .05$ ], indicating that the rating scores of current trials were assimilated to those of the 1-back trials. In addition, we also found that the magnitude of the sequential effect in the between-gender group was significantly smaller than that in the within-gender group [ $t(28) = 2.76$ ,  $p < .01$ ].

If judgments had a long-range response drift over the experimental session, a pseudosequential effect would be observed. However, the sequential effect for face attractiveness judgments occurred only for 1-back trials (Kondo et al., 2012). Also, the pseudosequential effect alone could not explain the difference of sequential effects between the same- and different-gender



**Fig. 1** Averaged regression coefficients for within- and between-gender groups in face attractiveness judgments (Exp. 1). The error bars indicate standard errors of the means

trials; since they were mixed randomly in the stimulus sequence, a response drift would have induced pseudosequential effects equally for both the same- and different-gender trials. Hence, we think a response drift is not responsible for the sequential effect observed in the present study.

### Experiment 2: Face roundness judgment

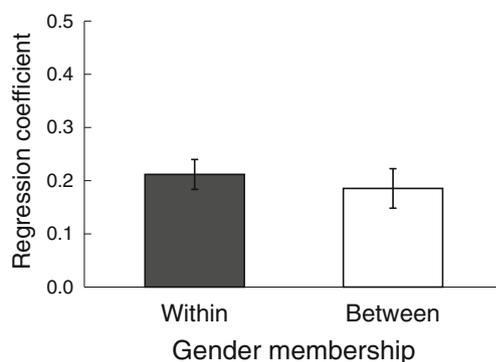
In Experiment 2, we examined whether gender membership would also affect sequential dependencies in judging a physical property (roundness) of the faces.

#### Method

A group of 29 paid volunteers were newly recruited (13 males and 16 females). The procedure was identical to that of Experiment 1, except that the participants were asked to rate the roundness of each face.

#### Results and discussion

Since the mean rating scores of roundness judgments for male and female faces were significantly different (male, mean = 3.53,  $SD = 0.53$ ; female, mean = 3.93,  $SD = 0.61$ ) [ $t(28) = -4.35, p < .0005$ ], we performed normalization as in Experiment 1, in order to estimate the sequential dependency without any biases. The regression coefficients were significantly greater than zero in both the between- and within-gender groups (Fig. 2) [within,  $t(28) = 7.41, p < .0001$ ; between,  $t(28) = 4.91, p < .0001$ ]; the judgment of face roundness was also assimilated to the previous judgments. However, in contrast to the face attractiveness judgment, the magnitude of the sequential effect in between-gender trials was not smaller than that in within-gender trials [ $t(28) = 0.61, n.s.$ ]. Thus, the gender membership seemed to play a modulatory role in the sequential effect for face attractiveness judgments, but not in that for face roundness judgments.



**Fig. 2** Averaged regression coefficients for within- and between-gender trials in face roundness judgments (Exp. 2). The error bars indicate standard errors of the means

### Experiment 3: Face intelligence judgment

In Experiment 3, we further investigated the influence of gender membership on the sequential effect in subjective judgments of faces other than attractiveness. For this purpose, we asked participants to judge how intelligent the person appeared to be.

#### Method

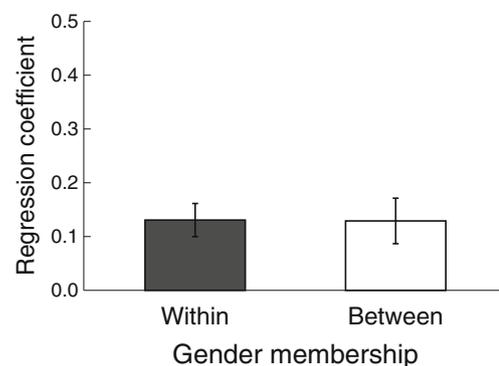
The procedure was identical to those of the previous experiments, except for the participants' task: They now rated how intelligent people looked from pictures of their faces. A group of 29 newly recruited participants (15 males and 14 females) took part in the experiment.

#### Results and discussion

The mean rating scores of intelligence judgments for male and female faces were not significantly different (male, mean = 4.11,  $SD = 0.47$ ; female, mean = 4.28,  $SD = 0.43$ ) [ $t(28) = -1.77, n.s.$ ]. However, for consistency, we performed normalization and submitted them to the regression analysis, as in the previous experiments. The regression coefficients were significantly greater than zero, irrespective of gender membership (Fig. 3) [within,  $t(28) = 4.17, p < .0005$ ; between,  $t(28) = 3.00, p < .01$ ]. However, the sequential dependencies of within- and between-gender trials were comparable [ $t(28) = 0.04, n.s.$ ]. This result suggests that not all subjective judgments of faces are influenced by gender membership.

#### Within-subjects correlations

The influence of gender membership on the sequential effect was significant only in the face attractiveness judgments (Exp. 1). However, the mean rating scores were also different between genders in Experiment 1 (as well as Exp. 2). Since we normalized the rating scores of male and female faces



**Fig. 3** Averaged regression coefficients for within- and between-gender trials in face intelligence judgments (Exp. 3). The error bars indicate standard errors of the means

separately, it was unlikely that the gender membership effect in Experiment 1 was an artifact in the regression analysis due to the difference in mean attractiveness. Even so, differences in the rating scores might lead to differential strategies of face categorization. For example, in an extreme case, the faces might receive either very high or very low ratings, independent of gender. To exclude this possibility, we conducted a within-subjects correlation analysis. If the difference of the rating scores in the male and female faces were responsible for the gender membership effect in Experiment 1, the larger were the differences in rating scores between male and female faces, the stronger the gender membership effect should be.

Figure 4 shows the difference in regression coefficients between the between- and within-gender trials (i.e., the strength of the gender membership effect) as a function of the absolute difference of the (nonnormalized) averaged rating scores between male and female faces. We found no correlation between them in any of the experiments [Exp. 1,  $r = .12$ ,  $t(27) = -0.65$ ; Exp. 2,  $r = .22$ ,  $t(27) = 1.16$ ; Exp. 3,  $r = -.11$ ,  $t(27) = -0.66$ , all  $ps > .25$ ]. This result ensured that it would be unlikely that the difference in face attractiveness between the male and female faces caused the gender membership effect in the sequential dependencies of face attractiveness.

## General discussion

The present study showed that the sequential effect of face evaluation was robust, irrespective of the attributes to be judged. The sequential judgments of faces were assimilated to the immediately preceding judgments for all attributes that we examined: attractiveness (Kondo et al., 2012), roundness, and intelligence. Furthermore, the gender membership of the face stimuli affected the sequential dependency only in the face attractiveness judgments, but not in the face intelligence or face roundness judgments.<sup>1</sup> Thus, the effect of gender membership category on sequential decision depended on the properties to be judged, even when the stimulus sets of faces to be judged were identical.

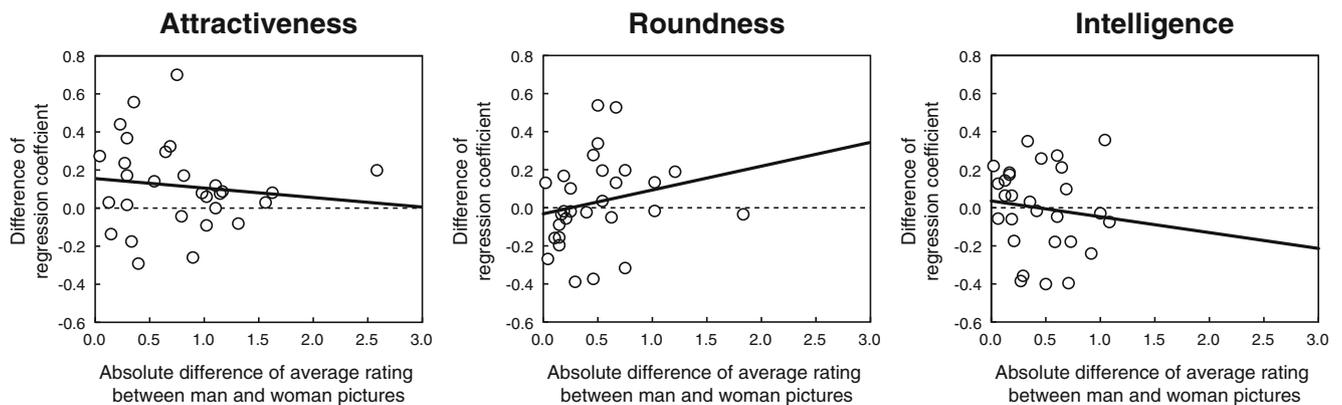
Sequential judgments are sometimes performed for a series of items from multiple categories and subcategories. We found that the sequential dependency in the face attractiveness

judgments was weaker when the genders of the faces in the current and 1-back trials were different (between-gender dependency) than when they were the same (within-gender dependency). These results are consistent with those of previous studies demonstrating that between-category sequential dependency was weaker than within-category dependency (Petzold & Haubensak, 2004; Ward, 1985). In the previous studies, however, the category of the sequentially presented stimuli was explicitly defined by stimulus modality (flash vs. sound; Ward, 1985) or by explicit instructions (Petzold & Haubensak, 2004). In contrast, we did not provide any explicit instruction about stimulus category. Nevertheless, the stimulus category (i.e., gender) automatically and selectively affected the attractiveness judgments of the faces. And, even when the stimulus sets were identical, the influence of gender membership on the sequential effects diminished in the face roundness and face intelligence judgments. These results suggest that the automatic modulation of sequential effects by stimulus category would be dependent on the task (i.e., the properties to be judged).

The specific influence of gender membership on sequential attractiveness judgments underlines that the gender of faces plays a special role in attractiveness judgments. Face attractiveness judgments might be performed using different reference frames for male and female faces. In the three experiments, we presented identical stimulus set of faces while varying the attributes to be judged. Face attractiveness, evaluated in Experiment 1, was a subjective value, whereas face roundness in Experiment 2 was a physical, objective attribute. However, the subjectivity of an attribute was not necessarily accompanied by gender membership effects, as, in Experiment 3, the gender membership had no influence on the sequential effect in the intelligence judgment, which was a rather subjective judgment. From an evolutionary perspective, face attractiveness judgments may be associated with mate selection (Thornhill & Gangestad, 1999). Gender is assumed to be a primal dimension of faces (Baudouin & Tiberghien, 2002), and gender would serve as a cue for forming category representations when the attractiveness of faces was of interest.

The reference points and/or facial features used in the sequential judgment of face attractiveness would be partially separated between male and female faces. In accordance with this idea, the visual features used for face recognition depend on the property to be judged (e.g., Schyns, Bonnar, & Gosselin, 2002). Moreover, Komori, Kawamura, and Ishihara (2009) analyzed the relation between face attractive and 80 facial features and showed that the attractiveness ratings of male and female faces were differently mapped onto the feature space. Thus, it is possible that differential reference points were formed for male and female faces. In the sequential effects, the most immediate events were used as a point of reference for the current judgments (Holland & Lockhead, 1968; Stewart,

<sup>1</sup> We compared the regression coefficients in paired  $t$  tests because this is a common practice in the research context of sequential effects. We also performed a test for the parallelism through an analysis of covariance for each participant. In Experiment 1, four out of 29 participants showed the significant positive effect (i.e., the slope in “same” trials was larger than that in “different” trials), and none showed a significant negative effect (i.e., the slope in “different” trials was larger than that in “same” trials). In Experiment 2, two were significantly positive, and one was negative. In Experiment 3, only one participant was significantly negative. Thus, although the statistical power might be very weak in this participant-based analysis, the results hinted at a tendency for the positive effect to be more dominant in Experiment 1.



**Fig. 4** Scatterplots of the differences in regression coefficients between within- and between-gender trials, as a function of the absolute difference in rating scores between male and female faces. Each data point indicates an individual's average

Brown, & Chater, 2005). The point of reference would move more easily within the enclosing reference frame, and hence the sequential dependency would be weakened if the reference frames used for the previous and current judgments were not identical. Our results implied that the reference frames for male and female faces were shared between face roundness and face intelligence judgments, whereas they were partially separated (e.g., dimorphic features; Komori et al. 2009) but partially overlapped (e.g., facial symmetry; Grammer & Thornhill, 1994; Perrett, Burt, Penton-Voak, & Lee, 1999) between genders in the face attractiveness judgments. Another possible reason for the specific influence of gender membership on sequential attractiveness judgments may be the gender of the raters. Some researchers have suggested that face attractiveness judgments depend on the gender of the raters (e.g., Foos & Clark, 2011; McLellan & McKelvie, 1993). However, for the data of Experiment 1, a  $4$  (gender sequence: male–male, female–female, male–female, female–male)  $\times 2$  (rater gender: male, female) analysis of variance showed no significant interaction [ $F(3, 81) = 0.44, p = .72$ ]. Although this could not be shown by the present data, it is still possible that the frame of reference might be different between when the stimuli have the same gender as the rater and when they do not.

Previous studies have shown that perceived attractiveness and perceived intelligence were correlated. For example, the correlations between attractiveness ratings and intelligence ratings in Zebrowitz, Hall, Murphy, and Rhodes (2002) were .51 (for adolescence) and .59 (for middle adulthood). Our study also showed a relatively strong correlation between the rating scores for attractiveness and intelligence ( $r = .65$ ). Therefore, the results are consistent with those of Zebrowitz et al. The correlation between attractiveness and intelligence scores suggests that visual features contribute to this overlap between judgments. However, as the overlap is not complete, other nonoverlapping features may contribute to the judgments, which may lead to the gender membership bias only for attractiveness judgments. It still remains to be investigated which features overlap and which do not.

Many studies have examined the effects of category membership on stimulus reproduction by using a serial reproduction task in which participants briefly viewed an object and then immediately estimated one of its attributes (e.g., the length of lines, size of shapes, or brightness of gray squares) from memory (e.g., Huttenlocher, Hedges, & Vevea, 2000; Sailor & Antoine, 2005). However, whether the categorical information of objects affects stimulus reproduction remains controversial. For example, Huttenlocher, Hedges, and Vevea argued that reproductions are biased toward the mean of each stimulus category, whereas Sailor and Antoine hypothesized that temporal contiguity (i.e., sequential effects) contributed to such a center bias and that the reconstruction process itself did not vary systematically as a function of category membership. Although the experimental procedures differed in multiple respects between the previous studies and the present study (subjective attractiveness judgments with faces), our findings might be considered as favoring the idea that local context established by the stimulus category of an immediately preceding trial would affect the subjective evaluation of a current stimulus. However, the specific conditions in which stimulus categories have influences remain to be explored in future studies.

In conclusion, our findings suggested that we cannot break free from the past in sequential decisions about faces; each judgment will bias the next judgment. In daily life, however, we have some occasions on which “fair” judgments of a person are required. In this case, the best that we can do is to know precisely how sequential dependencies will affect a specific situation. Face attractiveness explicitly and implicitly biases decisions (e.g., in job hiring; see Marlowe, Schneider, & Nelson, 1996). The results shown here warn us that we need to cautiously consider that our daily subjective judgments are accompanied by sequential dependencies, and that the magnitudes of the dependencies depend on the situation in which the judgment occurs—that is, on whether raters observe a within- or between-gender sequence.

**Author Note** This research was supported by Japan Society for the Promotion of Science to A.K. and Japan Science and Technology Agency (CREST) and Grant-in-Aid for Scientific Research (Grant No. 21300301) to K.W.

## References

- Baudouin, J. Y., & Tiberghien, G. (2002). Gender is a dimension of face recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 362–365.
- Damisch, L., Mussweiler, T., & Plessner, H. (2006). Olympic medals as fruits of comparison? Assimilation and contrast in sequential performance judgments. *Journal of Experimental Psychology: Applied*, *12*, 166.
- Foos, P. W., & Clark, M. C. (2011). Adult age and gender differences in perceptions of facial attractiveness: Beauty is in the eye of the older beholder. *Journal of Genetic Psychology: Research and Theory on Human Development*, *172*, 162–175.
- Grammer, K., & Thornhill, R. (1994). Human (*Homo sapiens*) facial attractiveness and sexual selection: The role of symmetry and averageness. *Journal of Comparative Psychology*, *108*, 233–242.
- Holland, M. K., & Lockhead, G. R. (1968). Sequential effects in absolute judgments of loudness. *Perception & Psychophysics*, *3*, 409–414. doi:10.3758/BF03205747
- Huttenlocher, J., Hedges, L. V., & Vevea, J. L. (2000). Why do categories affect stimulus judgment? *Journal of Experimental Psychology: General*, *129*, 220–241.
- Komori, M., Kawamura, S., & Ishihara, S. (2009). Effect of averageness and sexual dimorphism on the judgment of facial attractiveness. *Vision Research*, *49*, 862–869.
- Kondo, A., Takahashi, K., & Watanabe, K. (2012). Sequential effects in face-attractiveness judgment. *Perception*, *41*, 43–49.
- Marlowe, C. M., Schneider, S. L., & Nelson, C. E. (1996). Gender and attractiveness biases in hiring decisions: Are more experienced managers less biased. *Journal of Applied Psychology*, *81*, 11–21. American Psychological Association.
- Matthews, W. J., & Stewart, N. (2009). Psychophysics and the judgment of price: Judging complex objects on a non-physical dimension elicits sequential effects like those in perceptual tasks. *Judgment and Decision Making*, *4*, 64–81.
- McLellan, B., & McKelvie, S. J. (1993). Effects of age and gender on perceived facial attractiveness. *Canadian Journal of Behavioural Science*, *25*, 135–142.
- Page, L., & Page, K. (2010). Last shall be first: A field study of biases in sequential performance evaluation on the Idol series. *Journal of Economic Behavior and Organization*, *73*, 186–198.
- Perrett, D., Burt, D., Penton-Voak, I., & Lee, K. (1999). Symmetry and human facial attractiveness. *Evolution and Human Behavior*, *20*, 295–307.
- Petzold, P., & Haubensak, G. (2001). Higher order sequential effects in psychophysical judgments. *Perception & Psychophysics*, *63*, 969–978.
- Petzold, P., & Haubensak, G. (2004). The influence of category membership of stimuli on sequential effects in magnitude judgment. *Perception & Psychophysics*, *66*, 665–678. doi:10.3758/BF03194910
- Sailor, K. M., & Antoine, M. (2005). Is memory for stimulus magnitude Bayesian? *Memory & Cognition*, *33*, 840–851.
- Schyns, P. G., Bonnar, L., & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science*, *13*, 402–409.
- Stewart, N., Brown, G. D. A., & Chater, N. (2005). Absolute identification by relative judgment. *Psychological Review*, *112*, 881–911. doi:10.1037/0033-295X.112.4.881
- Thornhill, R., & Gangestad, S. (1999). Facial attractiveness. *Trends in Cognitive Sciences*, *3*, 452–460.
- Vlaev, I., & Chater, N. (2007). Context effects in games: Local versus global sequential effects on choice in the prisoner's dilemma game. *Judgment and Decision Making*, *2*, 380–389.
- Ward, L. M. (1985). Mixed-modality psychophysical scaling: Inter- and intramodality sequential dependencies as a function of lag. *Perception & Psychophysics*, *38*, 512–522. doi:10.3758/BF03207060
- Ward, L. M., & Lockhead, G. R. (1970). Sequential effect and memory in category judgment. *Journal of Experimental Psychology*, *84*, 27–34. doi:10.3758/BF03213031
- Zebrowitz, L. A., Hall, J. A., Murphy, N. A., & Rhodes, G. (2002). Looking smart and looking good: Facial cues to intelligence and their origins. *Personality and Social Psychology Bulletin*, *28*, 238–249.