Moreover, top-down characteristics (e.g., face familiarity) also affect the McGurk effect. Walker, Bruce, and O'Malley (1995) found that participants who are familiar with the face report less McGurk percepts than those who are unfamiliar with the face when the face and voice are from different persons. These studies imply that the McGurk effect can be modulated by either bottom-up or top-down attentional characteristics.

Value-driven attentional capture is a recently proposed mechanism of attention in addition to the salience-driven (bottom-up) and goal-driven (top-down) mechanisms (Anderson, 2013). Previous studies on value-driven attention, conducted in the visual domain, have shown that when stimuli are learned to predict reward, these stimuli would gain a competitive advantage that promotes attentional selection even when they are nonsalient and/or task irrelevant in perception (e.g., Anderson, 2013; Anderson, Laurent, & Yantis, 2011; Wang, Duan, Theeuwes, & Zhou, 2014; Wang, Yu, & Zhou, 2013). A few studies extended the concept to the cross-modal domain, showing that reward-associated sounds could affect the processing of visual stimuli. For example, auditory stimulus associated with high reward can increase the sensitivity of the perception of visual stimulus appearing simultaneously, even when sounds and reward associations are both irrelevant to the visual task (Pooresmaeili et al., 2014). Anderson (2016) demonstrated that relative to neutral sounds, previously reward-associated sounds capture attention, interfering more strongly with the performance of a visual task. However, it is currently unknown whether and how the value-driven mechanism of attention works in audiovisual speech perception in which the visual and auditory information is complex and highly relevant.

The face is the most important visual information in audiovisual speech perception. Raymond and O'Brien (2009) showed that a value-driven effect could be observed in the visual processing of faces. They trained participants to learn particular face–reward associations and then asked them to recognize whether a target face in an attentional blink (AB) task had been presented in the training phase. The authors found that the face recognition performance was higher for reward-associated faces compared with non-reward-associated faces. Moreover, while non-reward-associated-faces trials showed a typical AB effect, reward-associated-faces trials showed no AB effect, breaking through the constraints of AB on attentional selection. This study implies that value-associated faces would capture more attention and would be processed better than non-value-associated faces.

Considering that dynamic facial movements contain lots of visual information (e.g., mouth movements, other facial muscle movements, eye gaze), it is necessary to explore how people extract visual information from the dynamic talking faces for the purpose of audiovisual speech perception. By using the McGurk task and monitoring eye movements, previous studies have found that the mouth area of the talking face plays a critical role in the effect of visual information on audiovisual speech perception. In particular, perceivers show less time looking at the mouth area when the McGurk proportion decreases (i.e., when they make less use of visual information). For example, as the visual resolution of faces decreases, perceivers report fewer McGurk percepts and spend less time looking at the mouth area (Wilson, Alsius, Paré, & Munhall, 2016). Adding a concurrent cognitive task to the main McGurk task would decrease the McGurk proportion as well as the time looking at the mouth area (Buchan & Munhall, 2012). In addition, weak McGurk perceivers (i.e., perceivers who perceive the McGurk effect less frequently in general) fixate less on the talker's mouth area compared with strong McGurk perceivers (Gurler, Doyle, Walker, Magnotti, & Beauchamp, 2015; Hisanaga, Sekiyama, Igasaki, & Murayama, 2016).

The current study investigates whether and how the value-driven mechanism of attention works in audiovisual speech perception by using a training-test paradigm, which is often used in value-driven attention studies (e.g., Anderson, 2013, 2016; Anderson et al., 2011; Raymond & O'Brien, 2009; Wang et al., 2013). In the training phase, participants were asked to discriminate the gender of face pictures, in which correct responses to half of the faces could receive monetary rewards. In the test phase, participants were asked to identify the syllables that the talkers said in video clips (i.e., the McGurk task). Importantly, the talkers' faces had or had not been associated with reward in the previous training phase. In both phases, participants' eye movements were recorded with an eye tracker. Because value-associated faces could capture more attention and would be processed better than non-value-associated faces, we predicted that face–reward association would increase the influence of visual information on the audiovisual speech perception, resulting in higher McGurk proportion for reward-associated faces than for non-reward-associated faces. For the eye movement data, given the studies reviewed above, we predicted that participants would fixate longer on the oral area for reward-associated faces than for non-reward-associated faces.

Method

Participants

A group of 32 graduate or undergraduate students ranging in age from 18 to 26 years took part in the study for monetary compensation. They were all native speakers of Chinese and had the normal or corrected-to-normal vision and normal hearing; none of them reported a history of neurological or psychiatric disorders. This study was performed in accordance with the Declaration of Helsinki and was approved by the

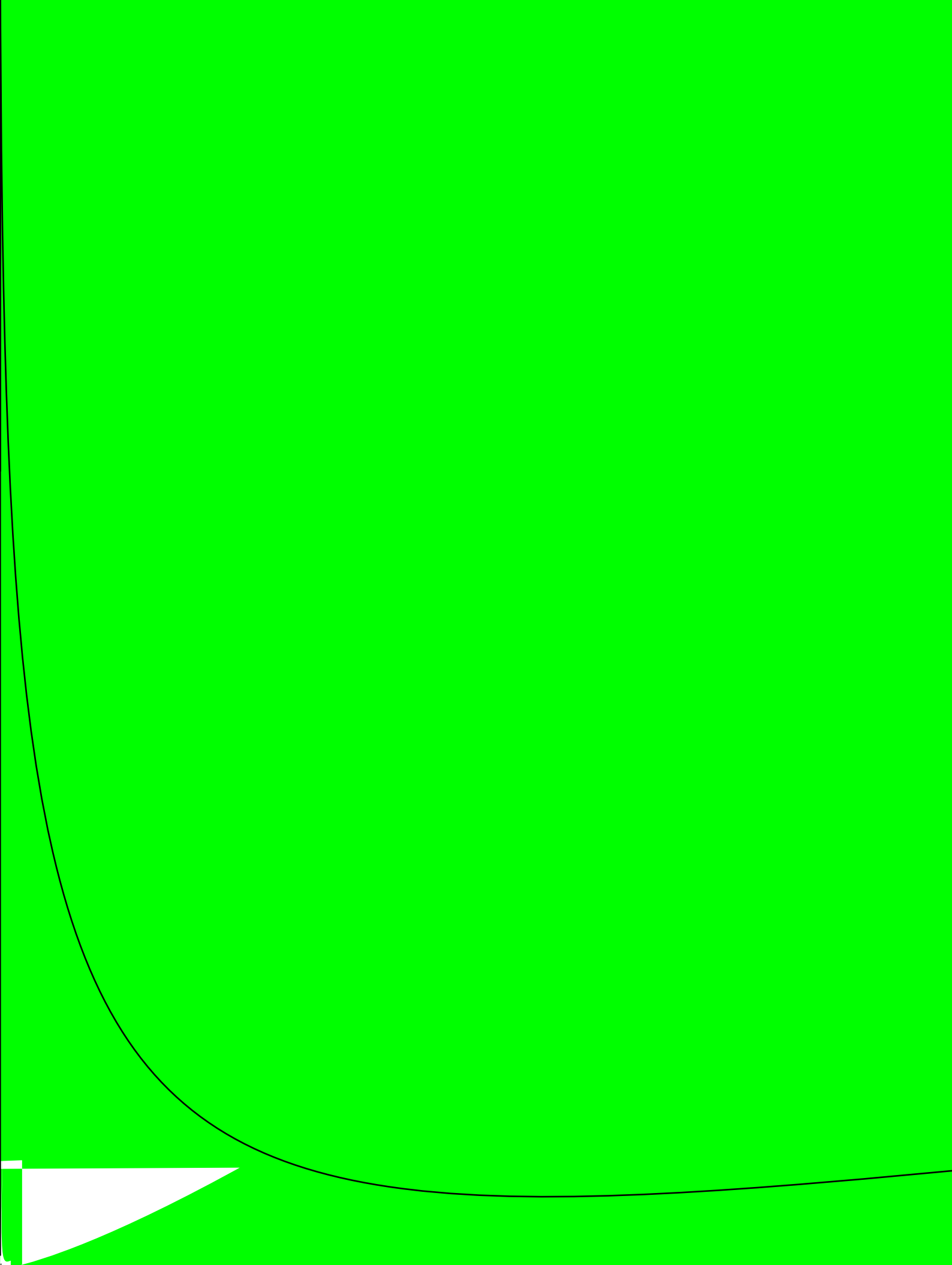
Committee for Protecting Human and Animal Subjects, School of Psychological and Cognitive Sciences, Peking University. Three participants were excluded, because they reported in the after-experiment interview that they concentrated on visual information in deciding the identity of spoken syllables during the task; another participant was excluded because of astigmatism that leads to poor quality of eye movement data. The remaining 28 participants were included in data analyses (17 females, mean age = 21.79, = 2.10). A power analysis was conducted by using G*Power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009; Faul, Erdfelder, Lang, & Buchner, 2007). Since we did not find a previous study that is similar to the current investigation, we referred to a study concerning the modulation of attention on the McGurk effect (Alsius et al., 2005). This study showed a moderate effect size (Experiment 1, Cohen's = 0.694). We estimated that we would need at least 19 participants, given Cohen's = 0.694, $\alpha = 0.05$, and power = 80%. In the present study, the number of participants (28) is higher than the suggested number (19).

Apparat s and materials

Visual stimuli were presented on a 17-inch SONY CRT monitor (refresh rate: 75 Hz, resolution: 1,024 × 768) connected to a DELL computer. Auditory stimuli were presented through an AKG headphone. The monitor was positioned 60 cm from the participant, and the head position was maintained using a chin rest. Eye tracking was performed using an EyeLink 1000 system. Stimulus presentation and participant's response recording were controlled by Psychophysics Toolbox (http://www.psychtoolbox.org/; Brainard, 1997) with MATLAB.

The audiovisual stimuli used in the test phase were developed based on eight original color video clips, which were recorded by an MI5 phone camera from two male and two female talkers (wearing white T-shirts) saying "ba" or "ga" without blinking. We edited the video and sound tracks with Windows Live video editing software and Cool Edit Pro 2.1, respectively. For each talker, three types of audiovisual stimuli were prepared by rematching sound and video recordings: congruent (visual and auditory matching, which included visual "ba" + auditory "ba," and visual "ga" + auditory "ga"), incongruent (visual and auditor8(ud)-17.61mi6(im)15.6999sm3(wa)16.899996.5(at)18.39999c76(sh)-376.10000682(m)13.1999998(sh)-22

me1h90-349.1000961(te)18.39999076(3o8-264.899993(3o8-363.1000tr5-4))12.1999ai5-40A)-3(rd)19.60000961(7e)11.19999t076(



and non-reward-associated faces. Specifically, a McGurk stimulus was regarded as a signal trial, and a filler stimulus (either the congruent or incongruent stimulus) was regarded as a noise trial. A response defined as a McGurk percept was regarded as a "yes" response, and a response not defined as a McGurk percept was regarded as a "no" response. Consequently, in the liberal definition of the McGurk effect, "hit" was defined as a response of any percept other than the auditory target to a McGurk stimulus, and "false alarm" was defined as a response of any percept other than the auditory target to a filler stimulus. By calculating the hit rate ( ) and false alarm rate ( ), we could obtain the [− (+) / 2] and the ′ (−) for each participant. Note that if or was 0 or 1, we would replace 0 with 0.5/ , and replace 1 with 1 − 0.5/ , where was the number of signal or noise trials (Stanislaw & Todorov, 1999). Similarly, in the conservative definition of the McGurk effect, "hit" was defined as a response of "da" to a McGurk stimulus, and "false alarm" was defined as a response of "da" to a filler stimulus.

Eye movement data

To explore how participants extract visual information from value-associated faces, we separately analyzed eye movement data in the training phase and in the test phase. These data were measured from the onset of the presentation of the face to the onset of gender discrimination response in the training phase and from the onset to the offset of each McGurk video clip in the test phase. Blinks, saccades, and fixation locations throughout each video clip were identified using the EyeLink Data Viewer. Interest areas (IAs) were created for each McGurk stimulus (see Fig. 2 for an example), including three rectangular bounding boxes for the left eye, the right eye, and the mouth of the talker; and two irregular bounding shapes for the nose/cheek and the forehead. All IAs covered the whole face, and the mouth IA was large enough to encompass the whole mouth, even at the maximal mouth opening. The critical IAs were essentially of the same size, allowing direct comparisons between different IAs. The IAs did not change during the display of video clips because the talkers' faces were relatively stationary.

Moreover, we analyzed the eye movement data over the time course of the presentation of the McGurk stimuli. We divided each video clip (1,500 ms in total) into three time intervals. The first time period covered 0 to 500 ms of the video clip in which the talkers kept stable without sound, and little visual information was provided in this period. The second time period covered 500 to 1,100 ms of the video clip in which the talkers opened their mouths and pronounced a syllable, with the sound played and most of visual information provided in this period. The third time period covered 1,100 to 1,500 ms of the video clip in which the talkers closed their mouths and returned to the original state without sound; little visual information was provided in this period.

We conducted 4 (IA: mouth vs. eyes vs. nose/cheek vs. forehead) × 3 (time period: first vs. second vs. third) × 2 (reward association: reward-associated vs. non-reward-associated) repeated-measures analysis of variance (ANOVA) on the proportion of looking time and the proportion of fixation number, respectively. In addition, for the purpose of illustration, we divided the whole McGurk stimulus presentation (1,500 ms) into 15 bins (100 ms for each) and calculated the proportion of looking time and fixation number on a particular IA in each bin. This was to further depict the change of the proportion of looking time and fixation number on different IAs over time, although we did not conduct statistical analyses for each bin.

Correlation analysis

Previous studies revealed that there were large individual differences in reporting McGurk percepts (Mallick, Magnotti, & Beauchamp, 2015) and in the eye movement pattern of looking at faces (Gurler et al., 2015). We expected that the reward-related differences of eye movement patterns in either the training or test phase (or both) might be related to the reward-related differences of McGurk proportion in the test phase. Thus, we explored the correlations between the eye movement data and the McGurk proportion. Specifically, we calculated the reward-related differences (i.e., subtracted the measures in the non-reward-associated condition from the measures in the reward-associated condition) of the looking time proportion and fixation number proportion in different interest areas for both the training and test phases; we also computed the reward-related changes of the McGurk proportion in the test phase. Then we conducted a series of correlation analyses between these measures.

Results

Reaction time and acc rac in the training phase

Participants identified the gender of reward-associated faces significantly faster than non-reward-associated faces (446 vs. 451 ms), (27) = −2.239, = .034, = 0.423, demonstrating that participants had learnt the face–reward association. There was no significant difference in terms of response accuracy between reward-associated and non-reward-associated faces (97.9% vs. 97.7%), (27) = 1.026, = .314.

McG rk effect in the test phase

The average accuracies in responding to the filler stimuli (i.e., congruent and incongruent stimuli) in different conditions were very high, ranging from 95.7% to 97.2%, indicating that participants performed the task carefully and effectively. For the McGurk stimuli, the proportion of each response category under each condition is shown in Table 1. According to the liberal definition of the McGurk percept (i.e., a response of any percept other than the auditory target was classified as a McGurk percept), the McGurk proportion was significantly higher for reward-associated faces than for non-reward-associated faces (60.1% vs. 49.9%), (27) = 2.438, = .022, = 0.461, which was consistent with our hypothesis.

According to the conservative definition of the McGurk percept (i.e., only a response of "da" was classified as a McGurk percept), the McGurk proportion was marginally higher for reward-associated faces than for non-reward-associated faces (52.2% vs. 43.6%), (27) = 1.788, = .085, = 0.338, which was consistent with the pattern reported above. In addition, the proportion of "ba" response (i.e., the true auditory target) was significantly lower for reward-associated faces than for non-reward-associated faces (39.9% vs. 50.1%), (27) = 2.438, = .022, = 0.461, mirroring the pattern for the liberally defined McGurk percept. The proportion of "ga" response (i.e., the true visual target) did not differ between reward-associated and non-reward-associated faces (2.7% vs. 2.3%), (27) = 0.350, = .729, = 0.066, nor did the proportion of "other" response (5.2% vs. 4.0%), (27) = 0.610, = .547, = 0.115. Note that there was a considerable variability in the McGurk proportion across participants based on either liberal or conservative definition (see Fig. 3), in line with a previous study (Mallick et al., 2015).

Table 1 Mean proportion of responses to McGurk stimuli with standard errors in parentheses

Reward association	Responses to McGurk stimuli (%)			
	"ba"	"ga"	"da"	"other"
ga				

Signal detection anal sis for the beha ioral data in the test phase

The signal detection analysis of the behavioral data based on the liberal definition of McGurk percept revealed that the was significantly lower for reward-associated faces than for non-reward-associated faces (0.735 vs. 0.998), (27) = 3.108, = .004, = 0.587, whereas the ′was significantly higher for reward-associated faces than for non-reward-associated faces (2.348 vs. 1.972), (27) = 2.089, = .046, = 0.395. This pattern was replicated in the signal detection analysis of behavioral data based on the conservative definition of McGurk percept, with the significantly lower for reward-associated faces than for non-reward-associated faces (1.018 vs. 1.247), (27) = 2.255, = .032, = 0.426, and the ′marginally higher for reward-associated faces than for non-reward-associated faces (2.218 vs. 1.882), (20) = 1.927, = .065, = 0.364.

E e mo ements in the training phase

We examined the proportion of looking time and fixation numbers on different interest areas (IAs) in the training phase. For the proportion of looking time, we conducted a 4 (IA: mouth vs. eyes vs. nose/cheek vs. forehead) × 2 (reward association: reward-associated vs. non-reward-associated) ANOVA, which showed only a significant main effect of IA, (3, 81) = 273.632, < .001, η_p^2 = .910, with the proportion of looking time on the nose/cheek IA (73.91%) significantly higher than the other three IAs (all s < .001) and the proportion on the eyes IA (23.81%) higher than on the mouth IA (1.03%) and the forehead IA (1.88%; all s < .001). Participants rarely looked at areas outside of the face. The number of fixations showed exactly the same pattern, with more fixations on the nose/cheek IA (73.25%) and the eye IA (23.87%) than on the mouth IA (1.16%) and the forehead IA (1.78%).

E e mo ements in the test phase: The proportion of looking time

Figure 4 illustrates the time course of the proportion of looking time (i.e., fixation time) at different IAs. The 4 (IA: mouth vs. eyes vs. nose/cheek vs. forehead) × 3 (time period: first vs. second vs. third) × 2 (reward association: reward-associated vs. non-reward-associated) ANOVA showed that the main effect of IA was significant, (3, 81) = 18.563, < .001, η_p^2 = .407. Planned comparisons showed that the proportion of looking time on the forehead IA was significantly lower than other three IAs (all s < .001), the proportion of

Fig. 3 Individual differences in the McGurk proportion based on the liberal definition of the McGurk percept. **a** Each participant's McGurk proportion for reward-associated and non-reward-associated faces. **b** The difference of the McGurk proportion between reward-associated and non-reward-associated faces for each participant

looking time on the mouth IA was marginally higher than on the eyes IA (= .073), and there were no differences between other IAs (all s > .269). The main effect of time period was also significant, (2, 54) = 126.548, < .001, η_p^2 = .824. Planned comparisons showed that the proportion of looking time on the first time period was significantly lower than the other two time periods (all s < .001), and there was no significant difference between the second and third periods (= .293). The main effect of reward association was not significant, (1, 27) = 0.040, = .843, η_p^2 = .001. The IA × Time Period interaction was significant, (6, 162) = 15.819, < .001, η_p^2 = .369, so was the IA × Reward Association interaction, (3, 81) = 2.897, = .040, η_p^2 = .097. The Time Period × Reward association interaction was not significant, (2, 54) = 0.558, = .576, η_p^2 = .020. Importantly, the three-way interaction between IA, time period, and reward association was significant, (6, 162) = 2.373, = .032, η_p^2 = .081, and we further explore this interaction below.

We conducted 4 (IA: mouth vs. eyes vs. nose/cheek vs. forehead) × 2 (reward association: reward-associated vs. non-reward-associated) repeated-measures ANOVA for the first, second, and third time periods, respectively. For the first time period (0 – 500 ms of the video clips; see Fig. 5, left panel), only the main effect of IA was significant, (3, 81) = 20.573, < .001, η_p^2 = .432. Planned comparisons showed that the proportion of looking time on the forehead IA was significantly lower than other three IAs (all s < .001), the proportion of looking time on the eyes IA was significantly lower than mouth IA (= .029) and nose/cheek IA (< .001), and there was no significant difference between mouth and nose/cheek IAs. The main effect of reward association and the interaction effect were not significant (all s > .493).

For the second time period (500–1,100 ms of the video clips; see Fig. 5, middle panel), the main effect of IA was significant, (3, 81) = 22.510, < .001, η_p^2 = .455. Planned comparisons showed that the proportion of looking time on the forehead IA was significantly lower than the other three IAs (all s < .001), the proportion of looking time on the mouth IA was significantly higher than the other three IAs (all s < .005), and there was no significant difference between eye and nose/cheek IA (= .277). The main effect of reward association was not significant, (1, 27) = 0.388, = .538, η_p^2 = .014. Importantly, the IA × Reward Association interaction was significant, (3, 81) = 2.908, = .040, η_p^2 = .097. Planned tests on simple effects showed that the proportion of looking time was significantly higher for reward-associated faces than for non-reward-associated faces (28.3% vs. 25.9%) on the nose/cheek IA, (27) = 2.328, = .028, = 0.440, although this effect did not reach significance if more stringent statistical tests were applied. This effect did not appear on

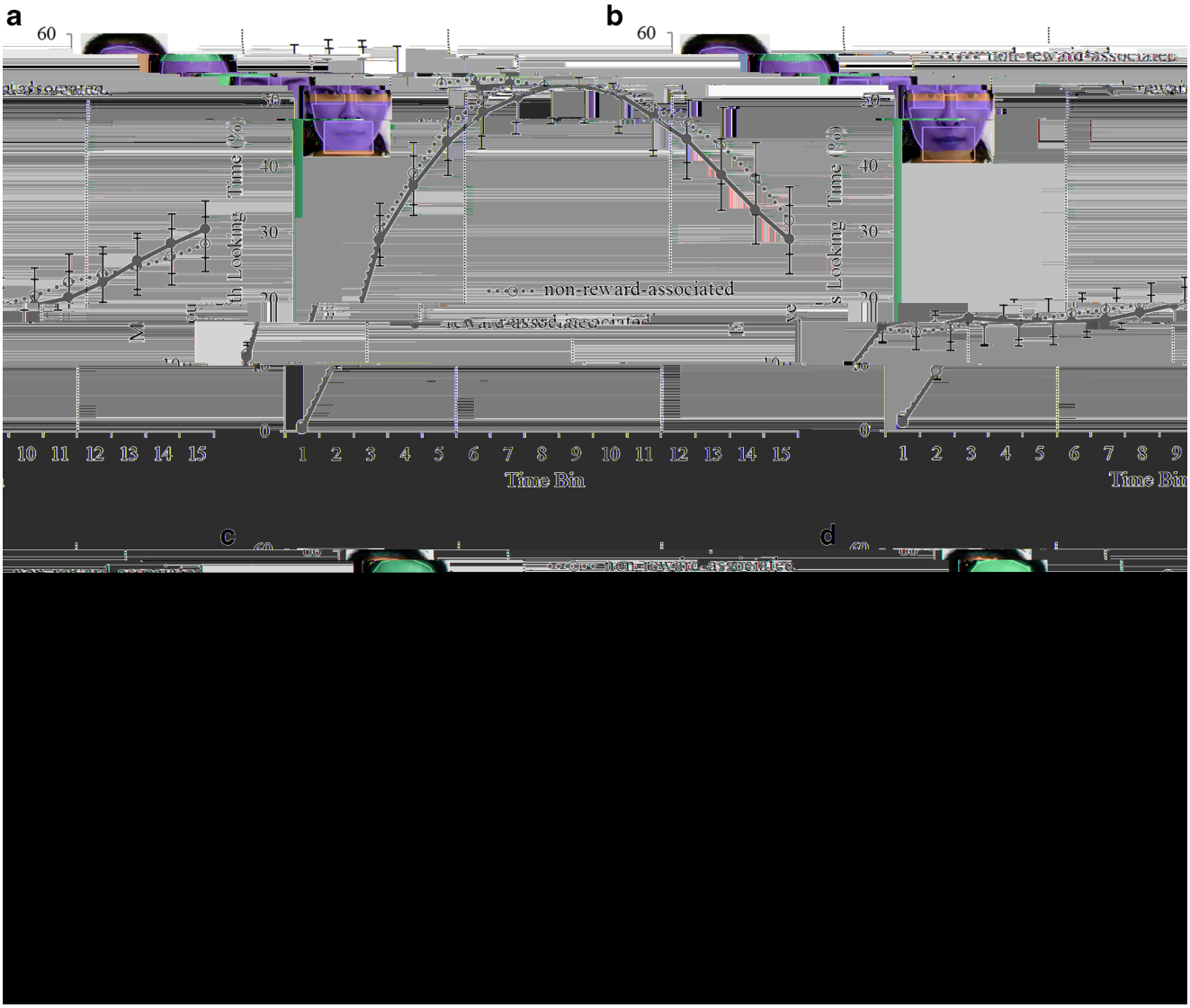


Fig. 4 Time course for the proportion of looking time on the interest area (IA) for (**a**) mouth, (**b**) eyes, (**c**) nose/cheek, and (**d**) forehead with standard errors. The whole McGurk stimulus presentation (1,500 ms) was divided into 15 time bins (100 ms for each) to further illustrate the change of the proportion of looking time on different IAs over time. The vertical lines separated time periods (i.e., 0–500 ms, 500–1,100 ms, and 1,100–1,500 ms of stimulus presentation) that we used in the statistical analyses

other IAs (all s > .101), suggesting that compared with non-reward-associated faces, participants looked at reward-associated faces longer, but only on the extraoral facial area, which is somewhat inconsistent with our original hypothesis.

For the third time period (1,100–1,500 ms of the video clips; see Fig. 5, right panel), the main effect of IA was significant, (3, 81) = 12.763, < .001, η_p^2 = .321. Planned comparisons showed that the proportion of looking time on the forehead IA was significantly lower than the other three IAs (all s < .001), and there were no significant differences between these three IAs (all s > .917). The main effect of reward association was not significant, (1, 27) = 0.241, = .627, η_p^2 = .009. But the IA × Reward Association interaction was significant, (3, 81) = 3.408, = .021, η_p^2 = .112. Planned tests on simple effects showed that the proportion of looking time was significantly lower for reward-associated faces than for non-reward-associated faces (36.0% vs. 40.0%) on the mouth IA, (27) = −2.122, = .043, = 0.401, although this effect would not survive if more stringent statistical tests were applied. There were not reward association effects on other IAs (all s > .098). The result here was surprising, as it indicated that participants were less likely to look at the mouth area of reward-associated faces, relatively to non-reward-associated faces, even though visual information in this area was thought to be a causer of McGurk effect. This is in contradictory to our original hypothesis.

We also collapsed data over the three time periods and conducted a 4 (IA: mouth vs. eyes vs. nose/cheek vs. forehead) × 2 (reward association: reward-associated vs. non-reward-associated) ANOVA. The IA × Reward Association

interaction was significant (see Fig. 6a). Planned tests on simple effects showed that the proportion of looking time on the nose/cheek IA was marginally higher for reward-associated faces than for non-reward-associated faces (26.9% vs. 25.1%), (27) = 2.034, = .052, = 0.384, although this effect would not survive when more stringent statistical tests were applied. The pattern here again demonstrated the importance of extraoral facial areas in the value-driven McGurk effect.

central target face in each trial, we analyzed the number of first fixation in a particular IA, excluding the fixations outside of the face. We found only a significant main effect of IA, (3, 81) = 17.038, < .001, η_p^2 = .387, with more fixations on the nose/cheek IA (36.66%) and the mouth IA (35.31%) than on the eye IA (19.55%) and the forehead IA (7.90%).

Correlation anal sis

Given that, in the training phase, participants spent the longest time looking at the nose/cheek area (73.91%) than any other areas, the correlation analysis was first conducted for this area. Over participants, the difference of the proportion of looking time at nose/cheek area between reward-associated and non-reward-associated faces in the training phase positively correlated with the difference of the McGurk proportion, either liberally or conservatively dent0821Tm[(Gi)14(,)-344.5bgetwee, the ditionk in the te phaserseg.y
,
tion offixationmber417.70000076(,)]TJ0TL/F111Tf9.96259975009.9625

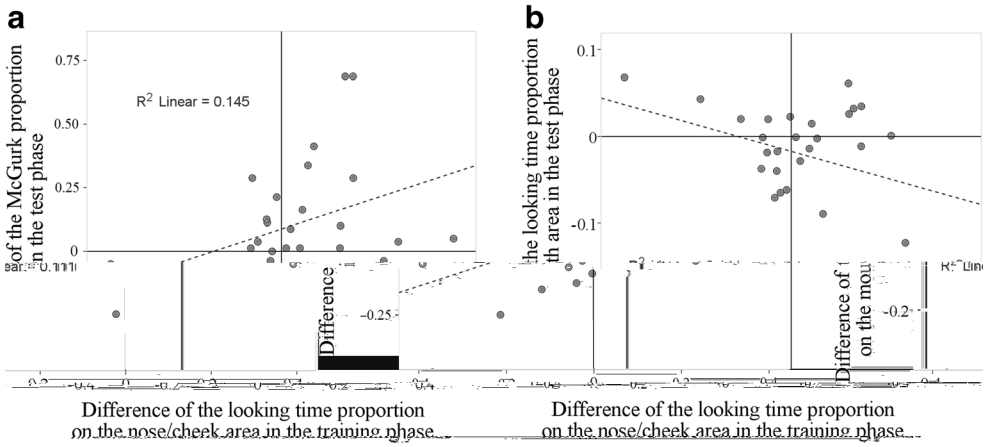


Fig. 8 Correlation analysis across the training and test phases. **a** The difference of the proportion of looking time at nose/cheek area between reward-associated and non-reward-associated faces in the training phase positively correlated with the difference of the McGurk proportion

test phase, however, the difference of the proportion of looking time at the nose/cheek area between reward-associated faces and non-reward-associated faces did not correlate with the difference of McGurk proportion, either liberally or conservatively defined, between the two conditions, = −.158, = .421; = −.071, = .719. The null effect was also observed for the proportion of fixation number, = −.179, = .362; = −.121, = .538. Similarly, in the test phase, the difference of the proportion of looking time at the mouth area between reward-associated and non-reward-associated faces did not correlate with the difference of McGurk proportion, either liberally or conservatively defined, between the two conditions, = .030, = .880; = −.090, = .648. The null effect was also observed for the proportion of fixation number, = .045, = .820; = −.042, = .831.

Discussion

The present study found that, in line with our prediction, participants would make more use of visual information for value-associated faces in audiovisual speech perception, with more McGurk percepts for reward-associated faces than for non-reward-associated faces. Value-associated stimuli have higher attentional priority (Anderson, 2013; Anderson et al., 2011). Participants devoted more attention to reward-associated faces than to non-reward-associated faces; the deeper processing of reward-associated faces increased the weight of visual information in audiovisual speech perception, resulting in more reports of the McGurk percepts. Convergent with this account, McGurk effect had been observed when less attention was assigned to the visual information (Alsius et al., 2005; Tiippana et al., 2004).

This account gains additional support from our signal detection analysis. Here, participants had lower for reward-associated faces than for non-reward-associated faces, demonstrating that participants tended to respond "yes" (i.e., having more liberal criterion) in audiovisual speech perception when faces were associated with value. The "yes" response, in the context of the present study, meant a response that was different from the actual auditory target. Thus, the change of here implies a tendency that participants made use of visual information even when the visual information was incongruent with the auditory information (Seilheimer, Rosenberg, & Angelaki, 2014).

Moreover, we speculate that multisensory integration processes may play a role in the value-driven McGurk effect. The signal detection analysis revealed that participants had higher ′ for reward-associated faces than for non-reward-associated faces, demonstrating that participants were more sensitive to the "signal" in audiovisual speech perception when faces were associated with value. The "signal" here refers to stimulus properties that could lead to the McGurk percept, and the change of sensitivity to these properties might be related to certain internal processes, such as multisensory integration for a McGurk stimulus. An fMRI study (Pooresmaeili et al., 2014) showed that reward associations modulate responses in multisensory processing regions (i.e., superior temporal sulcus [STS]) and other classical reward regions, but only the modulation strength of STS could predict the magnitude of the behavioral effect. The authors argued that multisensory regions may mediate the transfer of value signals across senses, rather than classical reward regions in the cross-modal context. Considering that our results demonstrate that the value-driven mechanism of attention works not only in simple cross-modal contexts (e.g., Anderson, 2016; Pooresmaeili et al., 2014) but also in the complex audiovisual

speech perception context, it is possible that multisensory integration processing was directly facilitated by reward association in the present study, resulting in more McGurk percepts for reward-associated faces. Nevertheless, it should be noticed that the McGurk effect cannot be equated with multisensory integration, because much more is involved with the McGurk effect than just multisensory integration, such as conflict resolution (e.g., Fernández et al., 2017; see also Alsius et al., 2018 for a review).

Furthermore, it should be mentioned that our results seem to show contrasts with a previous study in which Walker et al. (1995) investigated the influence of face familiarity (i.e., a form of value to some extent) on the McGurk effect, and found that participants who were familiar with the face reported McGurk percepts than those who were unfamiliar with the face when the face and voice were from different persons. However, there are key differences between the studies. Participants in our study did not know the talkers before, and all the talkers' faces in the training phase were static pictures and appeared at the same frequency. That is, participants had the same familiarity of all the talkers' static faces, and had no prior knowledge of the talkers' dynamic facial movements. Walker et al. (1995) defined the familiarity in terms of participants having had face-to-face interactions with the talker in daily life, which means that participants were familiar not only with the talkers' static faces but also with the talkers' dynamic facial movements and voices. As the authors mentioned, participants were able to use their prior knowledge of those familiar faces (expectations of what speech events were likely and of how these events were realized through dynamic facial movements); the incongruence between the visual and auditory modality was thus easier to be detected, resulting in less report of McGurk percepts. The authors also found that when the face and voice were from the same person, there were no differences in McGurk percepts between the participants who were familiar with the faces and the participants who were unfamiliar with them, a pattern recently replicated (Magnotti et al., 2018). It seems that there are fundamental differences between the value of familiarity and the value of reward, which merit future systematic studies.

Eye movements suggest how participants gather visual information from faces in audiovisual speech perception. In the present study, oral facial movements were the most important visual information in audiovisual speech perception, evidenced by the highest proportions of looking time and number of fixations on the mouth area than on other areas (see Fig. 4,5,6 and 7) and by the higher proportions of looking at the mouth IA in the audiovisual speech perception task (i.e., the test phase) compared with in the face recognition task (i.e., the training phase), in line with previous studies (e.g., Buchan & Munhall, 2012; Gurler et al., 2015; Wilson et al., 2016). One might predict that if McGurk proportion is increased for reward-associated faces, participants will look at the mouth area more often and/or with more time for reward-associated faces than for non-reward-associated faces. Surprisingly, however, we observed an opposite pattern in our results: Participants looked at the mouth area often and with time for reward-associated faces than for non-reward-associated faces, but they looked at the nose/cheek area often and with time for reward-associated faces than for non-reward-associated faces.

One possible ene ar

Munhall 2003), suggesting that information about mouth movements can be obtained from other areas in non-mouth-looking conditions.

Second, extraoral facial movements may provide useful visual information apart from the oral facial movements, which helps to elicit the McGurk effect. Thomas and Jordan (2004) manipulated the movements of the mouth and other facial areas independently, and found that the extraoral movements could promote the identification of audiovisual speech even when the mouth is kept static or removed from the face. Jordan and Thomas (2011) further found that the McGurk effect is observable even when the talker's face is occluded horizontally or diagonally (i.e., when the mouth area is occluded). In the present study, longer looking time and fixated more often on the extraoral area of reward-associated faces, compared with non-reward-associated faces, might help participants process the visual information provided by extraoral area, resulting in higher McGurk proportion.

To conclude, by associating faces with or without monetary reward in the training phase, we demonstrated that individuals could in the subsequent test phase report more McGurk percepts for reward-associated faces, relative to non-reward-associated faces, indicating that value-associated faces enhance the influence of visual information on audiovisual speech perception. The signal detection analysis revealed that participants have lower response criterion and higher sensory discriminability for reward-associated faces than for non-reward-associated faces, indicating that when the talking faces are associated with value, individuals tend to make more use of visual information in processing the McGurk stimuli. Surprisingly, we found that participants in the test phase had more looking time and number of fixations on the nose/cheek area of reward-associated faces than non-reward-associated faces; the opposite pattern was found for the mouth area. The correlation analysis revealed that the more participants looked at the nose/cheek area in the training phase due to reward, the more McGurk effect occurred in the test phase for reward-associated faces. These findings suggest that associating reward with a face may increase the atten-

stimuli, time, and response type. & ,
22