Computational exploration of task and attention modulation on holistic processing and left side bias effects in face recognition: the case of face drawing experts.

Bruno Galmar (brunogal@hku.hk)
Janet Hui-wen Hsiao (jhsiao@hku.hk)

Department of Psychology, University of Hong Kong
Pokfulam Road, Hong Kong SAR

Abstract

Drawing artists and non-drawers are like any adult both experts at face recognition. Yet, artists have a richer learning experience with faces: they were trained in rapid sketching of faces. Zhou, Cheng, Zhang and Wong (2011) found that drawing experts showed less holistic processing (HP) for face recognition than non-drawers. Using a computational model of face recognition that did not implement motor processing, we examined whether engagement of local attention and nature of the learning task could account for the reduced HP in drawers without the influence from motor experience. We showed that compared with the non-drawer model that had a global face input (i.e., Hsiao, Shieh & Cottrell, 2008), a drawer model that incorporated both global face and local facial parts (eyes and mouth) in the input showed reduced HP, suggesting the modulation of local attention engagement. In contrast, the other drawer model that used only global face input but learned to perform an additional face part identification task did not show the reduced HP effect. In addition, both drawer models demonstrated stronger left side (right hemisphere) bias than the non-drawer model. Our data thus suggest that engagement of local attention is sufficient to account for the reduced HP in drawers, and that HP and left side bias effects can be differentially modulated by visual attention or task requirements.

Keywords: Model of face recognition; Holistic processing; Hemispheric lateralization; Visual expertise.

Introduction

Visual expertise in subordinate-level discrimination has been extensively studied (e.g., Bukach, Gauthier, & Tarr, 2006), such as our expertise in recognizing individual faces. Several behavioral markers of visual expertise have been identified, including holistic processing (HP), which refers to the phenomenon of viewing faces as a whole instead of various parts (Bukach et al., 2006; although some argue that HP is specific to face recognition; e.g., McKone, Kanwisher, & Duchaine, 2007). Subsequent studies suggest a correlation between an increase in HP and expertise in subordinate-level individualization, as opposed to expertise in basic-level categorization (e.g., Wong, Palmeri, & Gauthier (2009)). For example, Wong et al. (2009) trained two participant groups to recognize an artificial object type (Ziggerins) with different training tasks: one group learned to rapidly individualize Ziggerins at the subordinate level, whereas the other group learned rapid sequential categorization at the basic level. The results showed that only the individuation experts showed an increase in HP, even though the two groups had the same amount of exposure to Ziggerins. This suggests that qualitatively different expertise processing can arise depending on the nature of the training task.

Such a qualitative difference of expertise processing resulting from different learning and training experience has been recently observed for face recognition. Zhou, Cheng, Zhang and Wong (2011) studied two groups: (a) an experimental group was composed of art students who had extensive formal training in sketching and drawing portraits, and (b) a control student group of non-drawers – i.e. who had no prior drawing background or education-. Hence, the two groups had different learning experience in processing faces. Non-drawers would show the typical face expertise any adult is endowed with: being able to recognize at least a thousand of faces. In contrast, art students would have internally assimilated an ordered procedure for rendering faces on a 2D surface (Balas & Sinha, 2007; Willenbrink & Willenbrink, 2012), for example: a) sketch the basic head proportion, b) sketch the overall head form and basic lines for features, c) place the brows and lips, and so on. Such a fine-grained procedure relies upon a mix of global and local processing, and featural and configural processing. Art students would not ignore face details which are critical to render a vivid portrait of an individual. Hence, art students are used to scrutinize a face and could be less engaged in HP than non-drawers. This educative guess is supported by eye-tracking studies (Miall & Tchalenko, 2001; Tchalenko et al. 2003) of eye movements of a skilled artist. Miall and Tchalenko (2001) proposed as an account of the visual encoding of the studied artist Ho: "*The capture of visual information detail by detail, rather than in a more holistic manner, is reflected in the way the drawing or painting is built up. Each detail and each element is of intrinsic importance.*" Using the complete composite paradigm of face recognition, Zhou et al. (2011) found less HP for art students than for non-drawers. Reduced HP with drawing expertise is not an isolate case. Previously, Hsiao and Cottrell (2009) found reduced HP for Chinese readers - who were experts at recognizing Chinese characters - compared with novice Chinese readers. Tso, Au, and Hsiao (2011) further showed that the reduction in HP found in expert Chinese readers depended on their writing rather than reading experience of Chinese characters, since proficient readers who had limited writing experience (i.e. Limited-writers) showed increased HP as compared with novices, in contrast to the reduced HP observed in Chinese readers who could read and write fluently (i.e., Writers; Tso, 2012).

In the present study, we aimed to examine the underlying mechanism accounting for the results in Zhou et al. (2011) through computational modeling and simulations. Computational modeling is an insightful tool to test ideas on the nature of cognition difficult to test with human subjects (McClelland, 2009). Motor experience, visual attention, and nature of the learning task are all potential factors that may account for drawers' reduced HP in face recognition. These factors may be difficult to disentangle within drawers so that the separate contribution of each to HP is not easily amenable to experimental study. Here, we aimed at testing two simplified models of drawing expertise that did not implement motor processing and to compare them with our previous model of face recognition (i.e., the intermediate convergence model in Hsiao, Shieh, & Cottrell, 2008), which is to serve as a non-drawer model, in order to examine whether visual attention and nature of the learning task can account for the reduced HP in drawers without the influence from motor experience. Through these two models, we postulated two hypotheses concerning how art students having developed expertise in the task of drawing faces could demonstrate reduced HP in face recognition compared with non-drawers.

The non-drawer model – called base model thereafter – shown in Figure 1 is trained to map face images to whole face identity. This global task is intended to reflect ordinary face recognition by non-drawers. The models of drawing expertise are not as purely global as the base model. They embed local processing in addition to the global face identification.

Rationale behind the first model of drawing expertise

Our first model of drawing expertise shown in Figure 2 is trained to map face, eyes and mouth images to whole face identity. Modeling the encoding of visual information from facial parts such as eyes and mouth to serve the task of whole face identity reflects the engagement by artists in local attention. Using eye-tracking, Tchalenko, Dempere-Marco, Hu, and Yang (2003) reported that artists do process individually facial parts and even scrutinize faces for informative details: "*[...] the experienced painter differed from the novice in his ability to repeatedly target saccades onto a small detail of the model's face, and to lock on to that detail in a steady fixation.*" Consistently, Zhou et al. (2011) showed that artists had slower response times (RT) compared with non-drawers. This could be because of the additional engagement of local attention on facial parts. The nature of this more local and prolonged visual engagement is translated in the first model of drawing expertise by a larger input layer compared with the base model. A drawing expert may manipulate more encoded visual inputs - as suggested by the expertise literature (Bransford, 2000) - but would still perform the same global identification task than the normal face recognizer. Because of the selective encoding of eyes and mouth in addition of global encoding of the face image, this model reflects engagement of both global and local attention at the encoding stage of visual processing.

Rationale behind the second model of drawing expertise

Our second model of drawing expertise shown in Figure 3 is trained to map face images to both whole face identity and cluster identities for mouth and eyes. Hence, the rationale is that artists use the same global attentional resources – i.e. the model has the same global input layer as the base model- but artists engage in a more analytical face recognition task. Here, given a face input, the model tries to recognize in addition to face identity, a mouth prototype (a kind of mouth) and a pair of eyes prototype (a kind of eyes). Such partitioning of eyes and mouth in kinds reflects that artists would engage in clustering facial features. This hypothesis is not only sound but also well-grounded. In his *Treatise on Painting*, the Renaissance genius Leornardo Da Vinci exposes some technical insights on how to develop the skills necessary to a portraitist (Rigaud, 1877). For example, in the section of "*How to remember the Form of a Face*", Da Vinci mentioned: "*If you wish to retain with facility the general look of a face, you must first learn how to draw well several faces, mouths, eyes, noses, chins, [...], all those principal parts which distinguish one man from another."* Then, we read: "*[...] noses are of ten different sorts: straight, bunched, concave, [...].*" In another section entitled "*Observations on drawing Portraits*", we read: "*The uniting of the nose with the brows is in two ways [...]. The forehead has three different forms.*"

Details on the implementation of these models are given in the next section. We trained the three models to either the same performance level in the whole face identification task or the same amount of epochs, and examined their difference in HP and lateralization. Face processing has been shown to involve right hemisphere (RH) lateralization, as indicated by the left side bias effect: a chimeric face made from two left half faces from the viewer's perspective is usually judged more similar to the original face than one made from two right half faces (Gilbert & Bakan, 1973). It is commonly assumed that HP is associated with RH lateralization. However, some experimental and computational studies (Hsiao & Cottrell, 2009; Hsiao & Cheung, 2011) showed the possibility of increased engagement of RH whereas decreased HP is measured. Another work on Chinese reading expertise (Tso, 2012) revealed a reduced HP for Chinese Writers as compared with Limited-writers; however there was no difference in left side bias between them. Our modeling work is hoped to also shed additional light on this issue.

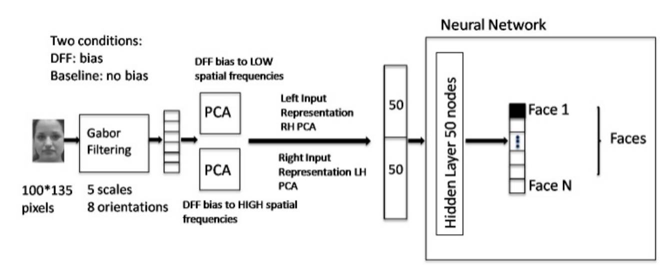

Figure 1: Base Model

Modeling Implementation

Base model for non-drawers

Face recognition by non-drawers is modeled by Hsiao et al.'s (2008) intermediate convergence model of face recognition. This model (Figure 1) incorporated several known observations about visual anatomy and neural computation. Hsiao et al.'s (2008) used Gabor responses over the input images to simulate neural responses of cells in the early visual area, and Principal Component Analysis (PCA) to simulate possible information extraction processes beyond the early visual area. They then used this PCA representation as the input to a two-layer neural network. In addition, they implemented a theory of hemispheric asymmetry in perception, Double Filtering by Frequency theory (DFF, Ivry & Robertson, 1997) in the model. The theory posits that visual information coming into the brain goes through two frequency-filtering stages. The first stage involves attentional selection of a task-relevant frequency range. At the second stage, the LH amplifies high spatial frequency (HSF) information, while the RH amplifies low spatial frequency (LSF) information. This differential frequency bias in the two hemispheres was implemented in the model by using two sigmoid functions assigning different weights to the Gabor responses in the two hemispheres. In the present implementation, the face input (100 x 135 pixels) was first filtered with a grid (6 x 6) of overlapping 2D Gabor filters in quadrature pairs at five scales and eight orientations. The five scales corresponded to 2 to 32 cycles per face (the task-relevant frequency range, depending on the image size. The maximum frequency should not exceed 2 pixels per cycle; the 6th scale, $2^6 = 64$ cycles per image exceeds the maximum frequency of the images, 100/2 = 50 cycles per image). The resulting Gabor vector representation of the face was split into left and right halves. The perceptual representation of each half was compressed into a 50-element representation. After PCA, each principal component was z-scored to equalize the contribution of each component in the model. The PCA representation was then fed to a feedforward network with one hidden layer of 50 nodes. The number of nodes was determined empirically to allow efficient training of the network of all the three models of the present study. The output layer of the neural network has one output for each of the 53 faces of the testing set. Face images were taken from the CAlifornia Facial Expressions dataset (CAFÉ; Dailey, Cottrell, & Reilly, 2001). We used two different neutral images for each face to constitute the training and testing sets. The neural network was trained with gradient descent with adaptive learning rate backpropagation from the MATLAB® Neural Network Toolbox (Version 7.0.3). All the networks were trained for both 400 epochs and 150 epochs. 400 epochs is enough for all the models to reach perfect recognition rates on the training sets and near perfect accuracy on testing sets. Training with only 150 epochs offers another viewpoint on the behavior of the three models by decreasing the ceiling effects observed with 400 epochs.

Implementation of model I of drawing expertise

Our first hypothesis states that drawing experts engage in local attention on specific facial features at the encoding stage in addition to the global encoding process shared with non-drawers. Hence, in addition to the face input, model I includes isolated mouth and isolated eyes as local inputs. We filtered mouth images (50 x 20 pixels) and eyes images (74 x 18 pixels) by a bank of Gabor filters of three scales and eight orientations. The three scales corresponded to 2 to 8 cycles per face (The maximum frequency should not exceed 2 pixels per cycle; the 4th scale, $2^4 = 16$ cycles per image exceeds the maximum frequency of the images, 18/2 = 9 cycles per image for eyes and 20/2 = 10 cycles per image for mouth). The size of the filtering grid (6 x 6) was the same for each kind of three - face, mouth and eyes - inputs reflecting the engagement of the same resources for processing the global face or anyone of the two local parts. The choice of eyes as a facial feature was motivated by Tchalenko et al.'s (2003) finding that artists primarily focused on eyes. We added also a bottom facial feature: mouth, richly informative for artists. After Gabor filtering, the vector representations of mouth and eyes followed the same scheme of splitting, weighting and compressing as the one for face input. Hence, the neural network of model I was fed with an input layer of length 300, with 100 PCA values for each of the three inputs. The model I of drawing expertise executes the same classification task as the base model. Hence, the two models have an identical output layer.

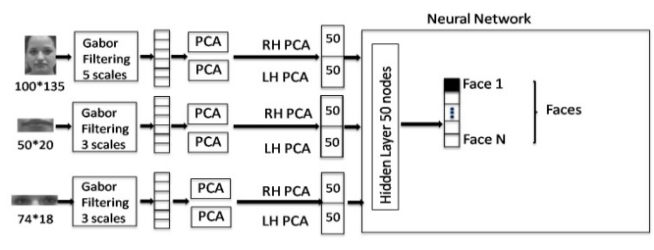

Figure 2: Model I of drawing expertise

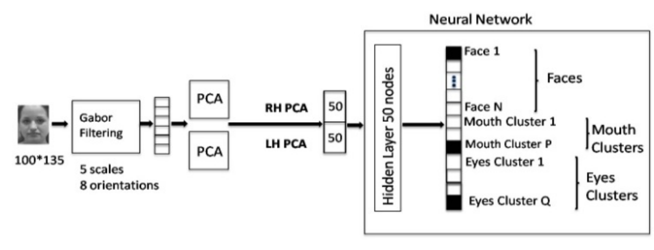

Figure 3: Model II of drawing expertise

Implementation of model II of drawing expertise

The second model of a drawing expert in Figure 3 is modified from the base model by adding at the classification stage of the neural network level two tasks. Namely, the model has to map the mouth and the eyes in the face input to respectively a "mouth cluster" and an "eyes cluster". This second model shares the same input layer with the base model. This means that both models use the same attentional or perceptual resources to encode the input face. However, the expert model is trained with a more analytic task than mere face identification. It has to perform a cluster mapping operation for mouth and eyes. Four eyes and four mouth clusters were defined based on a set of features for eyes and mouth mentioned in textbooks on drawing portraits. This clustering[1] yielded high recognition rates (> 98%) for mouth and eyes on both training and testing sets for both training durations.

Model of the composite task and measure of holistic processing

In human studies, HP is usually assessed through the composite paradigm (Young, Hellawell, & Hay, 1987). In this paradigm, two stimuli are presented briefly, either sequentially or simultaneously. Participants attend to either the top or bottom halves of the stimuli and judge whether they are the same or different. In congruent trials, the attended and irrelevant halves lead to the same response, whereas in incongruent trials, they lead to different responses. HP is indicated by interference from the irrelevant halves in matching the attended halves; it can be assessed by the performance difference between the congruent and the incongruent trials.

The holistic face processing effect has been accounted for by computational models. For example, Cottrell, Branson, and Calder (2002) trained a computational model to perform a face identification task and an expression judgment task, and showed that the model was able to account for HP effects in both tasks. Richler, Mach, Gauthier, and Palmeri (2007) also used a variant of Cottrell et al.'s (2002) model to account for the HP effect in face recognition. To assess HP in our three models, we applied the method used by Hsiao and Cheung (2011), which was derived from Richler et al. (2007). After training we attenuated the Gabor responses of either the top or bottom half of the images in the test set by multiplying a factor of 0.125 to simulate directing the models' attention to the bottom or top half of the images respectively. For the first model of drawing expertise, for mouth and eyes inputs, only the unattended part was attenuated (eyes are in the top half, mouth is in the bottom half; see Figure 5(a)). The complete composite design was used; it has been shown to be more robust than the partial composite paradigm (Richler, Cheung, & Gauthier, 2011). We created 4 types of stimulus pairs corresponding to the 4 conditions in Figure 4. Twenty pairs of images in each condition were randomly selected to form the materials (80 pairs in total). We calculated the correlation of the hidden layer representations in each pair as the similarity measure between them.

[1] We also considered using partitioning clustering methods such as k-means or PAM. However these methods yielded an optimal number of two clusters for eyes data. This result was not realistic from a human observer analysis. We finally preferred to keep the four eyes clusters found by human analysis.

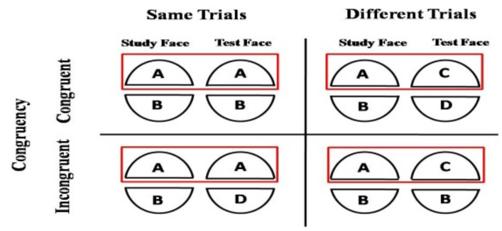

Figure 4: Design of the composite task, with top halves attended.

A threshold was set to be the midpoint between the mean correlation of the "same" stimulus pairs and that of the "different" stimulus pairs. We assumed that the model responded "same" when the correlation of a pair was higher than the threshold, and responded "different" when the correlation was lower than the threshold. The HP effect was indicated by the discrimination performance difference between the congruent and incongruent trials measured by d'.

Measuring hemispheric lateralization effect

The left side (RH) bias was assessed by the accuracy difference between recognizing a left-lateralized stimulus (carrying RH/LSF information) as the original stimulus and recognizing a right-lateralized stimulus (carrying LH/HSF information) as the original one. We defined RH lateralization (RH/LSF preference, Hsiao et al., 2008; Hsiao & Lam, in press) as the left side bias measured in the biased condition minus that measured in the baseline condition. For the first model of drawing expertise with additional mouth and eyes inputs, lateralized stimuli were also used following the scheme applied to the face input (see Figure 5 (b)).

Figure 5: (a) Illustrative example of a Congruent Same pair for the composite task where bottom half is attenuated. (b) Example of a left-lateralized stimulus for measuring lateralization effects. For (a) and (b), eyes and mouth parts were only used in Model I of drawing expertise.

Results

Model I of drawing expertise (Experiment 1)

As shown in Figure 6, the model I of expertise with an input layer completed with mouth and eyes local inputs demonstrated less HP than the base model after either 150 or 400 epochs of training. For the 400 epochs case (the perfect accuracy case on the training set), a directional t-test revealed that model I was statistically significantly less holistic than the base model, $t(798) = -1.76$, $p = .04$, confirming our hypothesis. The mean value of Δd' (Congruent d' – Incongruent d') for model I was smaller by a magnitude of 4 than the base model. This could be the result of a stronger ceiling effect. When decreasing the number of training epochs from 400 to 150, Δd' for model I was increased from 0.006 to 0.023, whereas Δd' for the base model increased from 0.026 to 0.063. Decreasing the number of epochs did not change the significantly lower amount of HP for model I compared to the base model, t(798) = -2.29, p = .011. Model I with its increased size of the input layer initially generalized better than the base model. For 150 epochs, model I outperformed the base model (98% versus 91% recognition rates on the testing sets). However, by 400 epochs, the base model caught up with model I, and both models had equally perfect recognition rates.

Concerning RH lateralization (see Figure 7), a t-test indicated that model I was significantly more subject to a left side bias than the base model, t(798) = 9, p < .001. For 150 epochs, the left side bias was further more accentuated for model I compared with the base model, t(798) = 16.03, p < .001.

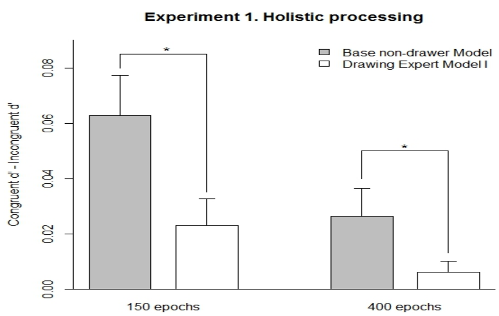

Figure 6: Experiment 1. Holistic Processing

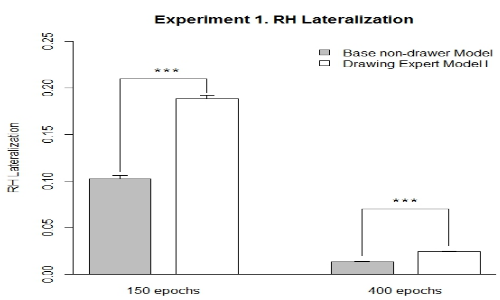

Figure 7: Experiment 1. RH Lateralization

Together the results indicated that our first model of drawing expertise compared with the base model of non-drawers is less holistic as measured by Δd' and is characterized by a stronger left side (RH) bias effect. This finding of more RH lateralization for the model of drawing expertise was somewhat unexpected: drawers by focusing on parts in addition to global processing could have engaged in more LH/HSF processing than non-drawers. However, the main result here is the replication of Zhou et al. (2011)'s finding of less HP for drawing experts compared with non-drawers.

Model II of drawing expertise (Experiment 2)

The model II of drawing expertise trained to recognize faces and to map mouths and eyes to respective clusters did not demonstrate less HP than the base model (see Figure 8). Statistical analysis showed that the expert model was as holistic as the base model for both 400 and 150 epochs, (t(798) = -0.38 , p = .35 ; t(798) = -1.12, p = .13). We expected model II to behave less holistically than the base model but it did not.

Concerning the left side (RH) bias, a t-test showed that model II was significantly more RH lateralized than the base model for both 400 and 150 epochs, (t(798) = 4.56, p < .001; t(798) = 3.17, p <.001). Again, this finding of more RH lateralization for the model of drawing expertise is somewhat unexpected: forcing the model to map eyes and mouth to cluster identities could have favored instead more LH/HSF processing (e.g., Hsiao & Lam, in press).

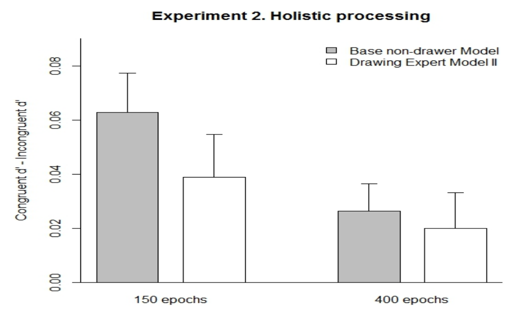

Figure 8: Experiment 2. Holistic processing

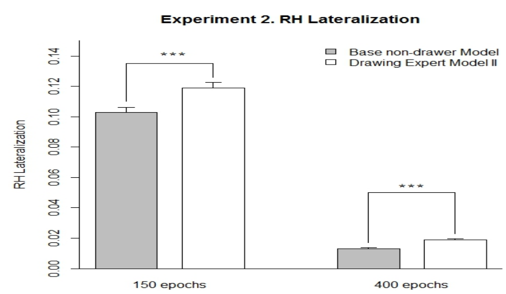

Figure 9: Experiment 2. RH Lateralization

Discussion & Conclusion

Through computational modeling, we explored the nature of drawing expertise and aimed at accounting for Zhou et al. (2011)'s finding of less HP for drawing experts compared to non-drawers. Our first model of drawing expertise relied on engagement of local attention on face parts at the encoding stage in addition to the mere global face encoding in the case of the base model. This model of drawing expertise was successful in accounting for a lesser amount of HP compared with the base model. In the second model of drawing expertise, we kept the input layer of the base model but added to the face identification task, a mapping task of eyes and mouth to cluster identities. This second model was as holistic as the base model. Our modeling idea of an enriched input layer of both local and global information for experts in model I is supported by eye-tracking studies

(Miall & Tchalenko, 2001; Tchalenko et al. 2003) of artists showing richer and more selective visual encoding by drawing experts compared with non-drawers.

Our findings of the two models of drawing expertise being more RH lateralized than the base model are congruent with the results of Hsiao and Cottrell (2009) on Chinese reading expertise. They found that Chinese character recognition experts have increased RH lateralization but reduced HP compared with novices. Like their results, our finding of increased RH lateralization but reduced HP for the first model of drawing expertise suggests that HP and RH lateralization may be separate processes that do not always go together, depending on the task requirement (Hsiao & Cheung, 2011). Our finding also provides a testable hypothesis that face drawers may exhibit stronger left side bias in face perception than non-drawers.

Tso (2012) showed that Chinese Writers and Limited-writers differed in HP but not in left side bias of Chinese characters. Drawers at first sight resemble Chinese Writers in that both achieved expertise through sharpening their motor and visual attention skills by eye-hand coordination while practicing their domain task. Nonetheless, the two groups may also differ in the following way. Chinese Writers were reinforced in a rote motor behavior while learning and copying the sequence of strokes for each character. However, drawers are not only challenged with each face's genuine and instantaneous uniqueness but critically have to render this uniqueness by capturing its gist in the details of the face. Hence, writing Chinese involves more rote motor learning than drawing faces; in contrast, drawers may develop better/finer visual attention skills than Chinese writers. Future work will examine whether our model can also account for Tso's (2012) finding in Chinese Writers and Limited-writers.

Our models of drawing expertise did not embed any motor component to represent motor drawing skills of experts. Hence, we showed that drawing experts and non-drawers could be sufficiently differentiated in terms of the nature (merely global versus both local and global) of attention during visual encoding of faces. We paved a first step in accounting for the nature of drawing expertise. It remains to be investigated what could be the contribution of motor expertise in drawing experts on the amount of HP they engage in.

Acknowledgments

We are grateful to the Research Grant Council of Hong Kong (project code HKU 745210H to J.H. Hsiao) and the HKU Seed Funding Program for Basic Research (project code 201011159124 to J.H. Hsiao).

References

Balas, B.J., & Sinha, P. (2007). Portraits and perception: configural information in creating and recognizing face images. *Spatial vision, 21*(1-2), 1-2.

Bransford, J. (2000). *How people learn: Brain, mind, experience, and school*: National Academies Press.

Bukach, C.M., Gauthier, I., & Tarr, M.J. (2006). Beyond faces and modularity: the power of an expertise framework. *Trends in cognitive sciences, 10*(4), 159-166.

Cottrell, G.W., Branson, K.M., & Calder, A.J. (2002). *Do expression and identity need separate representations.* Paper presented at the Proceedings of the 24th Annual Conference of the Cognitive Science Society.

Dailey, M., Cottrell, GW, & Reilly, J. (2001). California facial expressions (cafe). *unpublished digital images, University of California, San Diego*.

Gilbert, C., & Bakan, P. (1973). Visual asymmetry in perception of faces. *Neuropsychologia*, *11*, 355-362.

Hsiao, J.H., Shieh, D.X., & Cottrell, G.W. (2008). Convergence of the visual field split: Hemispheric modeling of face and object recognition. *Journal of Cognitive Neuroscience, 20*(12), 2298-2307.

Hsiao, J. H., & Cottrell, G. W. (2009). Not All Visual Expertise Is Holistic, but It May Be Leftist The Case of Chinese Character Recognition. *Psychological Science, 20*(4), 455-463.

Hsiao, J.H., & Cheung, K.C.F. (2011). Holistic processing and right hemisphere lateralization do not always go together—Evidence from computational modeling. *i-Perception, 2*(4), 242-242.

Hsiao, J.H., & Lam, S.M. (In press). The modulation of visual and task characteristics of a writing system on hemispheric lateralization in visual word recognition–A computational exploration. *Cognitive Science*.

Ivry, R.B., & Robertson, L.C. (1997). *The two sides of perception*: MIT Press.

McClelland, J.L. (2009). The place of modeling in cognitive science. *Topics in Cognitive Science, 1*(1), 11-38.

McKone, E., Kanwisher, N., & Duchaine, B.C. (2007). Can generic expertise explain special processing for faces? *Trends in cognitive sciences, 11*(1), 8-15.

Miall, R.C., & Tchalenko, J. (2001). A painter's eye movements: A study of eye and hand movement during portrait drawing. *Leonardo, 34*(1), 35-40.

Richler, J.J., Mack, M.L., Gauthier, I., & Palmeri, T.J. (2007). *Distinguishing between perceptual and decisional sources of holism in face processing.* Paper presented at the Proceesdings of the Twenty ninth Meeting of the Cognitive Science Society.

Richler, J. J., Cheung, O. S., & Gauthier, I. (2011). Holistic processing predicts face recognition. *Psychological Science, 22*(4), 464-471.

Rigaud, J.F. (1877). A Treatise on Painting by Leonardo da Vinci: London: George Bell.

Tchalenko, J., Dempere-Marco, L., Hu, XP, & Yang, GZ. (2003). Eye movement and voluntary control in portrait drawing. *Mind, 2*(3), 4.

Tso, R.V.Y., Au, T.K., & Hsiao, J.H. (2011). *The influence of writing experiences on holistic processing in Chinese character recognition.* Paper presented at the Proceedings of the 33rd Annual Conference of the Cognitive Science Society.

Tso, Ricky. (2012). *Motor experience modulates perceptual representation of objects: the case of Chinese character recognition.* (Postgraduate), The University of Hong Kong.

Willenbrink, M., & Willenbrink, M. (2012). *Drawing Portraits for the Absolute Beginner: A Clear & Easy Guide to Successful Portrait Drawing*: North Light Books.

Wong, A.C.N., Palmeri, T.J., & Gauthier, I. (2009). Conditions for Facelike Expertise With Objects Becoming a Ziggerin Expert—but Which Type? *Psychological Science, 20*(9), 1108-1117.

Young, A.W., Hellawell, D., & Hay, D.C. (1987). Configurational information in face perception. *Perception, 16*(6), 747-759.

Zhou, G., Cheng, Z., Zhang, X., & Wong, A. C. N. (2011). Smaller holistic processing of faces associated with face drawing experience. *Psychonomic bulletin & review*, 1-6.