SCIENTIFIC REP🞄RTS

OPEN

Disentangling locus of perceptual learning in the visual hierarchy of motion processing

Ruyuan Zhang [1,2] & Duje Tadin[1,3]

Visual perceptual learning (VPL) can lead to long-lasting perceptual improvements. One of the central topics in VPL studies is the locus of plasticity in the visual processing hierarchy. Here, we tackled this question in the context of motion processing. We took advantage of an established transition from component-dependent representations at the earliest level to pattern-dependent representations at the middle-level of cortical motion processing. Two groups of participants were trained on the same motion direction identification task using either grating or plaid stimuli. A set of pre- and post-training tests was used to determine the degree of learning specificity and generalizability. This approach allowed us to disentangle contributions from different levels of processing stages to behavioral improvements. We observed a complete bi-directional transfer of learning between component and pattern stimuli that moved to the same directions, indicating learning-induced plasticity associated with intermediate levels of motion processing. Moreover, we found that motion VPL is specific to the trained stimulus direction, speed, size, and contrast, diminishing the possibility of non-sensory decision-level enhancements. Taken together, these results indicate that, at least for the type of stimuli and the task used here, motion VPL most likely alters visual computation associated with signals at the middle stage of motion processing.

A large body of evidence has shown that the human visual system can gain long-lasting perceptual improvements following several sessions of perceptual training. This phenomenon, called visual perceptual learning (VPL), has been an active area of research because VPL is a remarkable demonstration that human vision can remain plastic even in adulthood[1,2]. Numerous studies have revealed training-induced perceptual improvements on a wide range of visual tasks, including low-level contrast and orientation discrimination tasks[3–6], mid-level motion and form tasks[7–9] and even high-level object and face recognition tasks[10,11].

While the robustness of learning effects is well established, debate persists with respect to the mechanisms underlying VPL. Early psychophysical work found that learning effects are usually confined to the trained parameters[6,12]. Such strong specificity suggests that VPL most likely takes place within low-level visual areas (e.g., V1 or V2) since neurons therein exhibit narrow ranges of spatial and feature selectivity (e.g., orientation, motion direction). Recent evidence, however, challenges this idea by revealing an increasing number of cases where the transfer of VPL is viable to novel stimulus conditions and tasks[13,14]. This is consistent with an involvement of higher-level visual areas, wherein neurons usually respond to larger spatial areas and more complex stimulus features. Some studies even suggest the contributions from the brain areas that process non-sensory attributes. For instance, perceptual learning might manifest as the change of decision variables encoded in the prefrontal cortex[15]. Alternatively, perceptual learning might facilitate encoding of abstract concepts representing basic visual features (e.g., orientation and contrast)[16] or lead to a better set of task-specific rules[17]. Given that these theories postulate changes beyond canonical sensory mechanisms, we refer to them as 'non-sensory' learning processes.

The task of linking VPL to specific brain areas is complicated by the complex functional specializations of the brain[18]. The brain includes multiple brain regions that are organized into a coarse, but richly interconnected hierarchy[19,20]. Even a simple perceptual choice likely arises from the interplay among multiple brain regions. One strategy is to take advantage of visual processes where links between behavior and neural structures are

[1]Department of Brain & Cognitive Sciences and Center for Visual Science, University of Rochester, Rochester, NY, 14627, USA. [2]Center for Magnetic Resonance Research, Department of Radiology, University of Minnesota at Twin Cities, Minneapolis, MN, 55455, USA. [3]Departments of Ophthalmology and Neuroscience, University of Rochester Medical Center, Rochester, NY, 14642, USA. Correspondence and requests for materials should be addressed to R.Z. (email: ruyuanzhang@gmail.com)
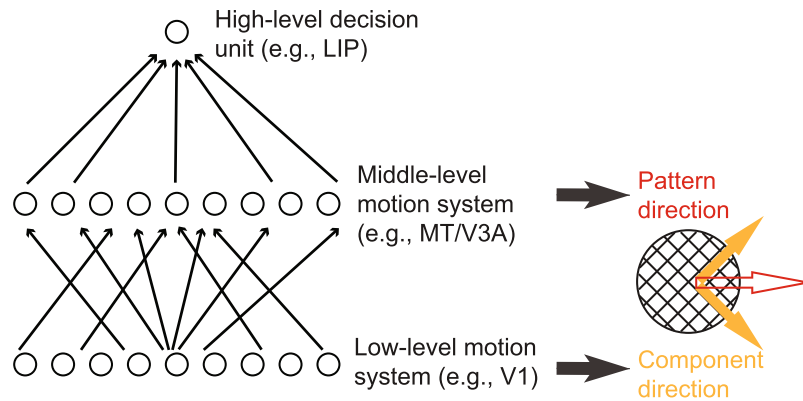
Figure 1. A simplified hierarchy of visual motion processing with three hierarchical stages[21]. Neurons in the low-level motion system respond best to component directions of plaid stimuli (full orange arrows). Neurons in the middle-level motion system respond selectively to the perceived pattern motion direction (empty red arrow). In this and all subsequent figures, empty arrows indicate faster perceived speed associated with plaid motion. The third stage involves complex sensory and non-sensory high-level cognitive processes.

well established. Here, we focus on VPL in context of motion perception, a perceptual domain where we have a relatively advanced understanding of different processing stages[21]. In primates, neurons selective to motion direction first occur in the earliest cortical areas V1 and V2[22]. However, conscious motion perception is most closely linked to intermediate visual areas, such as MT and V3A. These areas contain a large portion of neurons showing strong preferential responses to different motion directions[23–26]. In addition, perceptual decisions based on motion stimuli have been linked to several higher-level brain areas (e.g., lateral intraparietal cortex (LIP) and prefrontal cortex). These areas are often described as "evidence accumulators" that integrate sensory information provided by the upstream motion processing units in order to form perceptual decisions and guide visual behaviors[27,28] (but see ref.[29]). Finally, non-sensory attributes, such as task rules and decision strategies, encoded in high-level cognitive areas, can also mediate performance in motion perception tasks[30]. This complex hierarchy can be operationalized as a symbolic three-layer network (Fig. 1). This network consists of a low-level (e.g., V1/V2), a middle-level (e.g., MT/V3A) and a high-level (e.g., LIP, prefrontal cortex) processing stage.

In contrast with the established understanding of visual motion processing stages, less is known about their role in VPL. A prominent piece of work comes form Law and Gold[28] who showed that, in non-human primates, perceptual training alters the neural signals in LIP rather than MT. On the other hand, Pilly and colleagues[31] found that task-irrelevant VPL with random-dot motion does not transfer if only contrast polarity of the motion signal changes. This result is consistent with the properties of low-level motion processing, and is consistent with earlier evidence in favor of low-level motion VPL[32]. More recent work, however, showed transfer of learning between first- and second-order motion[16]. Given evidence for distinct mechanisms for first- and second-order motion[33,34], the authors interpreted their result as indicating high-level conceptual learning that is independent of key stimulus properties. However, there is also evidence for a common representation of first- and second-order motion direction across the wide range of visual areas[35]—evidence that would predict generalization of VPL between these two types of motion. In sum, the nature of motion VPL remains an open question.

Here, we aimed to provide new insights into motion VPL by taking advantage of an established transition from component-dependent representations at the earliest level to pattern-dependent representations in the middle-level of cortical motion processing[36–38]. A plaid stimulus composed of two obliquely moving gratings (Fig. 1) is generally perceived as a rigid object moving horizontally[39]. While typical MT neurons faithfully respond to the perceived motion direction in moving plaids[38], neurons in V1 primarily respond to the directions of two component gratings[36,37]. This dissociation allows us to psychophysically infer the main locus of signal changes associated with motion VPL. If learning effects are specific to the trained component motion, irrespective of the perceived pattern motion, it would indicate component-dependent learning that is accompanied with neural signal changes predominantly at the lowest levels of motion processing. Conversely, if learning effects are specific to the perceived pattern motion, it would point toward pattern-dependent learning associated with signals at middle-levels of motion processing. If this learning still retains some sensory-level specificity (e.g., speed, size and velocity), we can conclude that non-sensory cognitive processes are not the major drivers of the observed behavioral improvement.

Methods

Participants and apparatus. Fourteen undergraduate students from University of Rochester (18 to 22 years old, 5 males and 9 females) took part in this study. All participants had normal or corrected-to-normal vision. The Research Subjects Review Board at the University of Rochester approved experimental protocols and confirmed that all research was performed in accordance with relevant guidelines and regulations. Informed consent forms were obtained from all participants. Stimuli were generated by Matlab Psychtoolbox[40] and presented using customized digital light processing (DLP) projector (DepthQ WXGA 360 driven by a NVIDIA Quadro FX 4800 at 1280 × 720 resolution). The projector frame rate was 360 Hz, resulting in discrete 2.78-ms frames. DLP
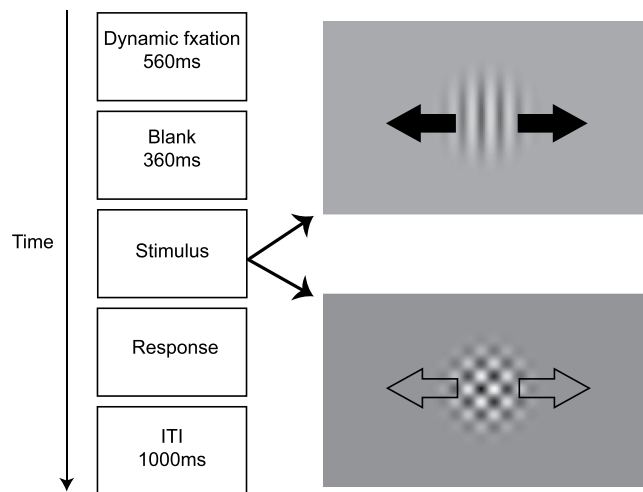
Figure 2. Task illustration showing the trial structure used for all training, pre- and post-test conditions. Participants viewed a moving stimulus that was either a grating or a plaid (arrows are for illustration purposes only). Stimulus duration varied on each trial, as determined by two interleaved staircases. Participants indicated the perceived stimulus direction via button press (e.g., left vs. right in this case).

projectors are natively linear, and this was verified with a Minolta LS-110 photometer. Viewing distance was 61.5 inches, with a projected image of 46.74 × 25 inches.

Stimulus and task settings. Participants were randomly assigned into two groups – one group trained on component motion (grating; N = 8) and another group trained on pattern motion stimuli (plaid; N = 6). All participants were tested and trained on a two-alternative forced choice motion direction identification task (Fig. 2), reporting the perceived stimulus motion direction via key press. Auditory feedback was provided after each trial during the training phase but not at pre-/post-test (to minimize learning effects in pre-/post-test). To facilitate fixation, we used the following fixation sequence (Fig. 2): a fixation circle (0.8° radius) appeared after each key press response and, the circle shrank to 0.13° over 200 ms, remained at that size for 360 ms, and then disappeared 360 ms before stimulus onset. We found in our previous work that this dynamic fixation sequence was very effective in guiding eye gaze to the center of the screen before the stimulus onset[41]. The inter-trial interval was 1000 ms.

As detailed below, the two training groups used partially overlapping sets of pre- and post-test conditions. We selected this design to limit pre- and post-test sessions to only the most diagnostic test conditions for each group. This allowed us to test the bidirectional transfer between component and pattern motion, as well as the dependency of learning transfer on several key low-level stimulus features.

In the component-training group, the training stimulus was a grating (contrast = 50%, drift speed = 4°/s, radius = 8°, 2D raised cosine spatial envelope; spatial frequency = 1 cycle/°; Fig. 3Aa). Training motion directions were either left/right or up/down, counterbalanced across participants. Motion directions for other stimulus conditions were adjusted according to the directions of trained stimuli. During the pre- and post-test, temporal duration thresholds (defined by the full-width at half-height of a hybrid between a Gaussian and a trapezoidal temporal envelope; see[42] for details) were measured across another five stimulus conditions: (1) a plaid stimulus moving in the trained directions (Fig. 3Ab); (2) a plaid stimulus containing the trained component and moving 45° away from the trained direction (Fig. 3Ac); (3–5) moving gratings that matched the trained grating except that they differed in (3) direction and orientation (orthogonal to the trained direction; Fig. 3Ad), (4) stimulus size (radius = 1°, Fig. 3Ae), and (5) contrast (contrast = 2%, Fig. 3Af). Note that all plaid stimuli in this study were additive plaids, following previous electrophysiological studies that demonstrated the transition from component-selective and pattern-selective neurons along the visual hierarchy of motion processing[36,43,44]. For all plaids, motion directions of two components were always orthogonal and deviated ±45° away from the plaid motion direction (empty arrows in Fig. 3).

For the plaid-training group, the plaid stimuli consisted of two orthogonal component gratings (component contrast = 50%, size = 8°, 2D raised cosine spatial envelope; component spatial frequency = 1 cycle/°; Fig. 3Ba). Component drift speed was 4°/s, which resulted in the plaid velocity of 5.66°/s. Training motion directions were either left/right or up/down, counterbalanced across participants. In addition to the trained condition, duration thresholds were measured for five additional pre- and post-test stimulus conditions: (1) a moving grating with the same apparent direction and speed as the trained plaid stimulus (Fig. 3Bb); (2, 3) two component gratings that constituted the trained plaid stimulus (i.e., gratings with direction ±45° away from the trained directions, Fig. 3Bc; Note that these data were collected in two separate blocks, each testing one motion direction axis, and subsequently averaged to get a single threshold estimate); (4) a plaid stimulus moving to the untrained directions, but comprised of same static component features (Fig. 3Bd); (5) a grating moving in the trained directions (left/right) but with the original plaid apparent speed (speed = 5.66°/s, Fig. 3Be).
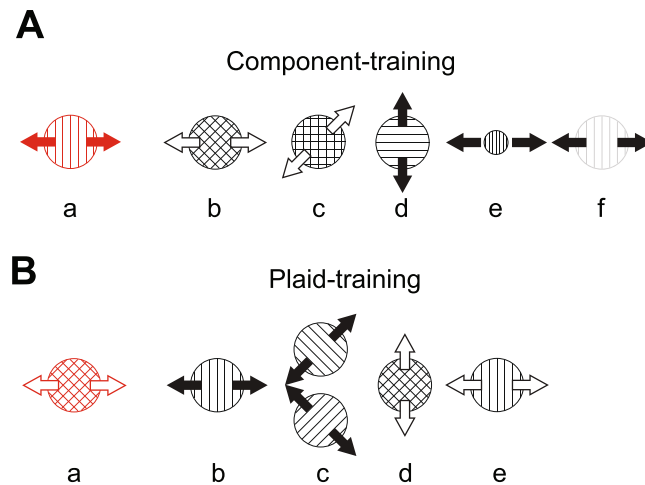
Figure 3. Pre- and post-training stimuli for (**A**) the component-training group and (**B**) the plaid-training group. The red icons show the training stimulus for each group. These conventions are kept in subsequent figures. With the exception of stimulus *e* in panel B, the speed of all grating stimuli was 4°/s (marked by solid arrows). The plaid component speed was also 4°/s, which resulted in the apparent plaid speed of 5.66°/s (marked by empty arrows). To assess the effect of stimulus speed on transfer of learning, we also included a grating whose speed matched the plaid speed of 5.66°/s (stimulus *e* in panel B, empty arrows). Although all stimulus conditions were conducted together, we analyze and present data into two batches: bidirectional transfer between component and pattern motion (Fig. 5) and transfer to other stimulus features (Fig. 6).
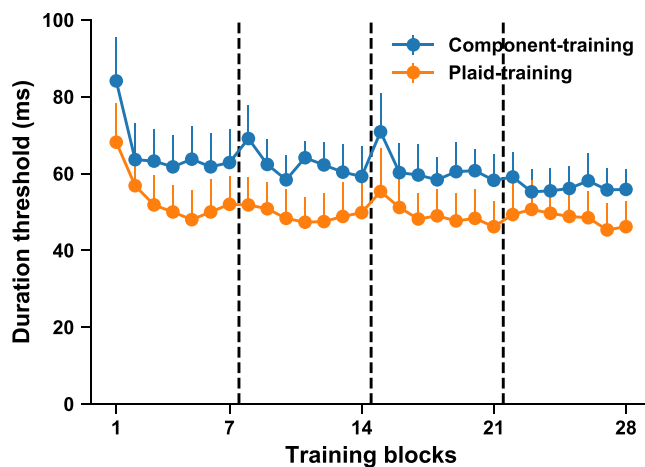

Figure 4. Learning curves for the component- (blue) and the plaid-training (orange) groups. This color scheme is kept in subsequent figures. Data are thresholds for 28 training blocks, tested over 4 days of training. Learning curves for individual subjects are shown in Fig. S1. Vertical dashed lines separate data for four training days. Note that the plaid-training group showed overall lower duration thresholds. This is expected given the faster apparent speed of plaid stimuli and known effects of stimulus speed on temporal duration thresholds[65]. Error bars are SEM across subjects.

Experimental procedures and data analysis. Pre- and post-test consisted of six randomly ordered blocks corresponding to different stimulus conditions (the trained stimulus, plus 5 additional stimulus conditions, as described above). In each block, stimulus durations were controlled by two 80-trial interleaved staircases (a 2-down-1-up staircase and a 3-down-1-up staircase), yielding 160 trials for each threshold estimate. The initial starting durations for two staircases were 100 ms and 110 ms, respectively. Pre- and post-test measurements were conducted on day 2 and day 7, respectively. On day 1, each participant completed a practice phase that was identical to the pre- and post-test battery, except that each block consisted of only 60 trials. The purpose of this practice day was to help stabilize pre-test measurements. The perceptual training lasted four days (days 3–6). On each day, participants completed seven 100-trial blocks, resulting in a total of 28 training blocks. For the first training block on the first training day (day 3), the initial starting durations for the two staircases were 100 ms and 110 ms. For all subsequent training blocks, the initial stimulus durations were the durations in the final trials of the two staircases in the previous training block. All participants completed these seven experimental sessions within 14 days (Fig. 4).
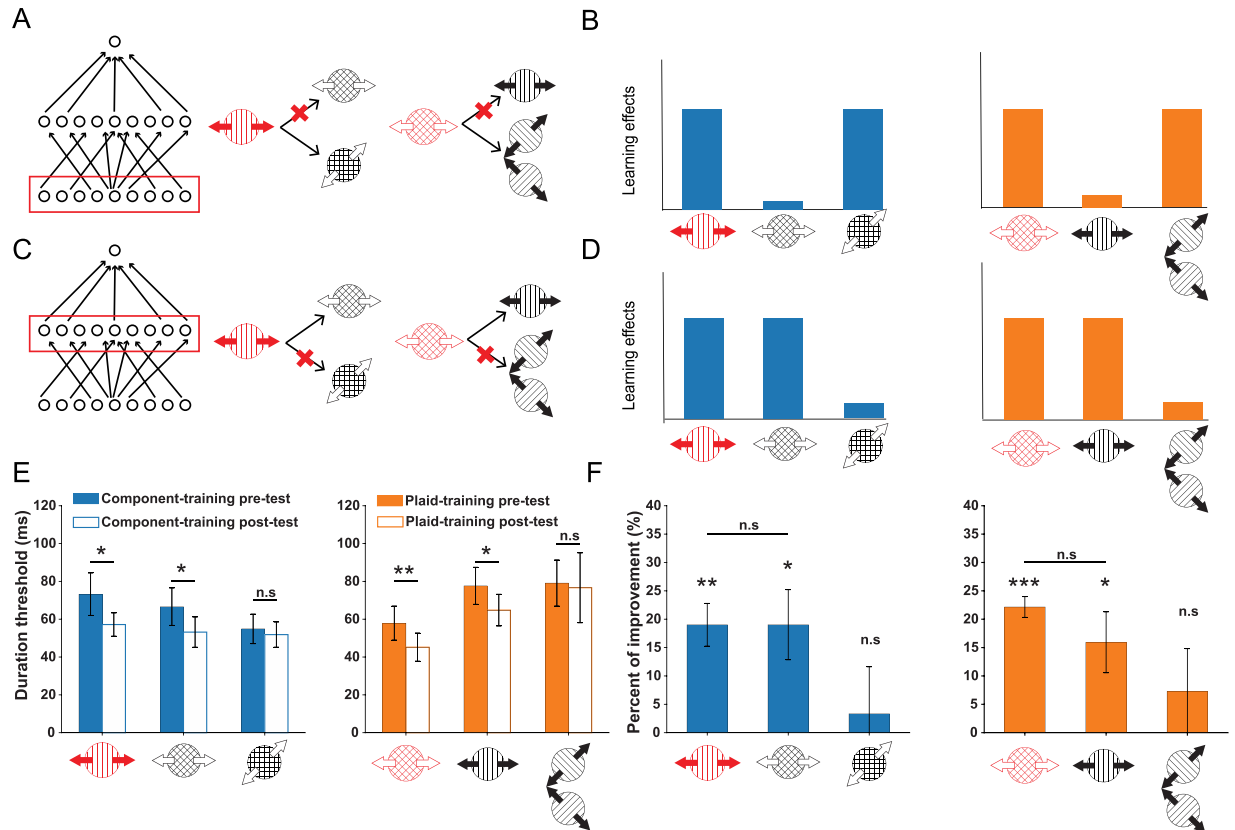
Figure 5. Schematic illustrations (**A,C**), predictions (**B,D**) and empirical results (**E,F**) for component-dependent and pattern-dependent VPL. **A**. The component-dependent VPL takes place at the lowest level of motion processing, as indicated by the red rectangle. Here, training on a component stimulus should only transfer to the plaid stimulus that comprises the trained component. Moreover, training on a plaid stimulus should only transfer to its two constituent components. **B**. Learning effects as predicted by component-dependent learning in panel A. **C,D**. Illustrations of the pattern-dependent perceptual learning and its predicted learning effects, following conventions in panels A and B. Here, plasticity involves the middle stage of motion processing. **E**. Duration thresholds at pre-/post-test across stimulus conditions in the component (left panel) and the plaid training (right panel), respectively. **F**. Learning effects quantified as percent of improvement (PI%) across stimulus conditions and training regimes. The overall pattern mimics the predictions in panel D, indicating plasticity associated with the middle-level of motion analysis. For all subplots, error bars denote ± 1 SEM across subjects. Significance symbol conventions are *p < 0.05; **p < 0.01; ***p < 0.001; n.s.: non-significant. Same definitions of error bars and symbol conventions are kept for all figures in this paper.

To estimate duration thresholds for each pre- and post-test condition, we fit Weibull psychometric functions to 160 trials of raw data using the maximum likelihood method, estimating the thresholds at 82% correct. The amount of learning in each condition was estimated by computing percent of improvement (PI):

$$PI = \frac{threshold_{pre} - threshold_{post}}{threshold_{pre}} * 100\%$$

(1)

where $threshold_{pre}$ and $threshold_{post}$ indicate duration thresholds for the corresponding pre- and post-test stimulus conditions. We used paired t-tests for comparisons of pre- and post-test thresholds and for comparison of PI across stimulus conditions. One-sample t-tests were used for assessing the statistical significance of PI against the null hypothesis of 0% PI. All t-tests were two-tailed and performed using Matlab Statistical and Machine Learning Toolbox.

Results

Effective perceptual learning for both component and plaid stimuli.
We first examined whether our training procedure was sufficient to result in perceptual improvement. Here, for each group, we compared pre- and post-test thresholds for the trained stimulus condition. The results revealed significant improvements in thresholds for both the component- and the plaid-training group (Fig. 5E; t(7) = 2.79, p = 0.0268 and t(5) = 6.28, p = 0.0015, respectively). We also computed percent of improvement (PI, see Equation 1), and found significantly positive PIs for both groups (Fig. 5F; t(7) = 5.06, p = 0.0015; t(5) = 12.04, p = 6.97 × 10^−6), with each group showing about a 20% improvement in performance.
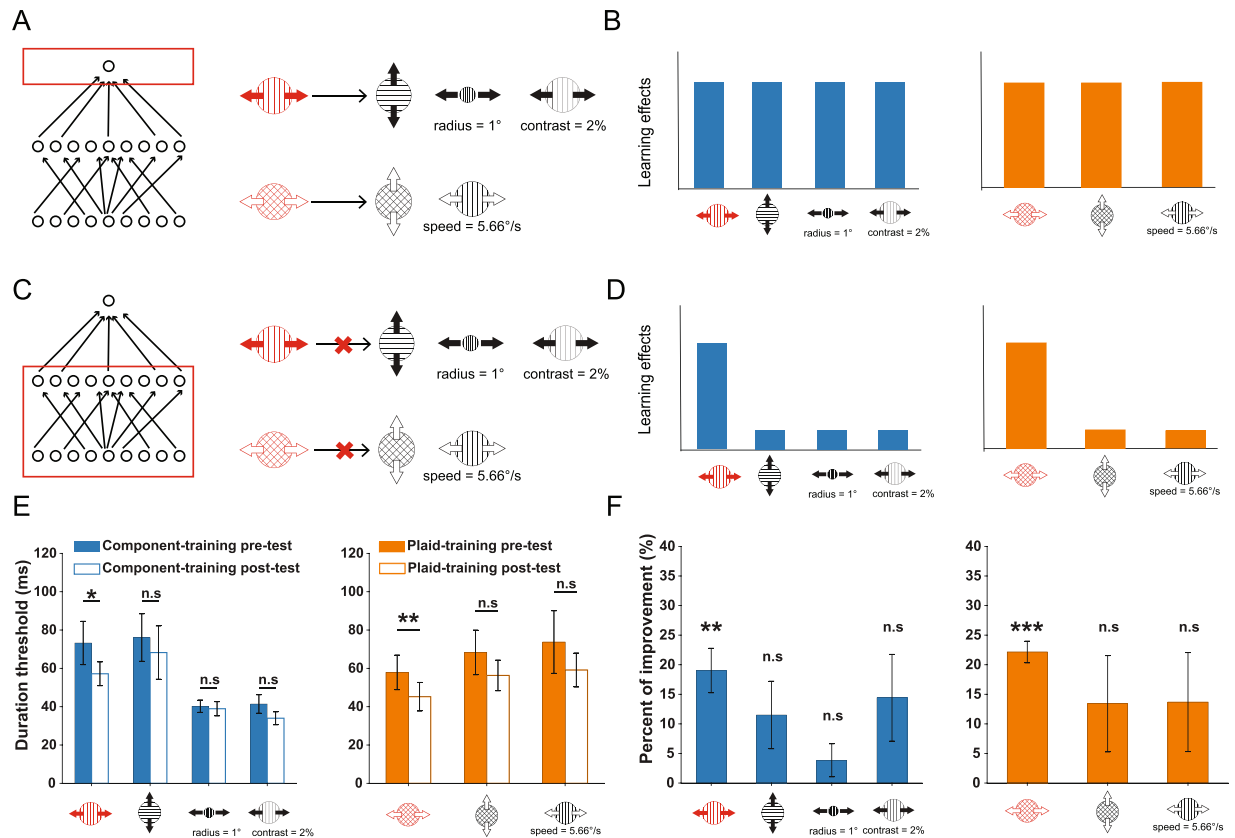
Figure 6. Specificity of motion VPL to basic sensory features. *A,B*. Illustrations and predictions of the plasticity at the highest-level stage (e.g., prefontal cortex) in the three-layer network. This mechanism predicts that training on a component or a plaid stimulus should be generalizable regardless of the variations in low-level visual features, such as direction, speed, size, and contrast. *C,D*. Illustrations and predictions akin to panels A and B, expect that the plasticity occurs within the general sensory representation stage. This scheme predicts that training on a component or a plaid stimulus should exhibit minimal transfer to the stimuli that differ in basic visual features. *E*. Duration thresholds at pre-/post-test across stimulus conditions in two training groups. *F*. Empirical learning effects, quantified as percent of improvement (PI%), across stimuli and training groups. The transfer pattern of learning is more consistent with predictions in panel D. No significant transfer in all other stimulus conditions is noted, implying the plasticity associated with the sensory representation level as shown in panel C.

Bidirectional transfer of learning between component and plaid motions. The main focus of this paper is to examine the transfer of perceptual learning to a range of diagnostic stimulus conditions. A two-stage criterion was used to assess transfer of learning. First, we concluded that learning transfers to a stimulus condition if the pre-/post-test difference on this condition was statistically significant. If a stimulus condition passed this first test, then we compared its PI to the corresponding trained condition (i.e., either trained component or trained plaid). If the transfer PI was significantly smaller than the trained PI, the result was described as a "partial transfer". Alternatively, if the PI for a transfer condition was not statistically smaller than the PI for its corresponding trained condition, we referred to it as "complete transfer", according to an established convention in VPL research[13,16,17].

The key aim of this study was to determine whether perceptual training leads to changes associated within low-level component-dependent motion processing or changes reflecting middle-level pattern-dependent motion processing. To be precise, component-dependent learning predicts that training on a component motion stimulus should only transfer to the plaid composed of the trained component gratings, and that training on a plaid stimulus should only transfer to its two constituent components (Fig. 5A,B). On the contrary, pattern-dependent learning predicts that training on a component motion stimulus or on a plaid motion stimulus that moves in the same directions should mutually transfer to each other (Fig. 5C,D).

Our results were consistent with learning in pattern-dependent mechanisms. First, perceptual training on a component grating significantly reduced the duration thresholds on the plaid that moved in the same apparent direction as the trained grating (Fig. 5E left panel, pre-/post-test, t(7) = 2.88, p = 0.0237; Fig. 5F left panel, PI, t(7) = 3.08, p = 0.0178). More importantly, the PI was statistically equivalent to the PI on the trained grating (Fig. 5F left panel; t(7) = 0.002, p = 0.999). Consistently, perceptual training on a plaid stimulus also transferred to the component grating that moved in the same apparent direction (Fig. 5E right panel, pre-/post-test, t(5) = 3.336, p = 0.0207; Fig. 5F right panel, PI, t(5) = 2.971, p = 0.0311). Also, the PI on the untrained component was not statistically different from the PI on the original trained plaid (Fig. 5F right panel, PI, t(5) = 1.29,

p = 0.2533). The bidirectional transfer between the component and the plaid stimuli that moved to the same directions suggest that perceptual training most likely alters the computation in the visual units that process the pattern motion direction. Moreover, training effects on a component did not significantly transfer to a plaid that included the trained component, but moved in a different direction (Fig. 5E left panel, pre-/post-test, t(7) = 0.784, p = 0.4586; Fig. 5F left panel, PI, t(7) = 0.405, p = 0.6978). Plaid training also did not improve the performance on its two constituent components (Fig. 5E right panel, pre-/post-test, t(5) = 0.305, p = 0.7709; Fig. 5F right panel, PI, t(5) = 0.963, p = 0.3797). Taken together, these findings suggest that motion VPL, at least for the type of stimuli and the task used here, is more closely associated with pattern-dependent middle-level mechanisms than the changes involving low-level motion mechanisms.

Specificities to direction, speed, size, and contrast.

We have thus far focused on experimentally disentangling component-dependent from pattern-dependent VPL, with the results arguing against low-level component-dependent VPL. What remains unclear, however, is whether the perceptual training led to enhancements in the processing of sensory features or high-level non-sensory attributes. For instance, participants might learn motion directions as abstract concepts[16] or be more familiar with the general task statistics (e.g., stimulus timing, stimulus-response association[17]). In this case, plasticity takes place in higher brain hierarchy that is independent of the sensory processing. To further delineate the learning-induced changes in the non-sensory (Fig. 6A,B) or the sensory processing (Fig. 6C,D), we examined the tolerance of our training across several other forms of stimulus variations, i.e., direction, speed, size, and contrast. The prediction is that if the plasticity is largely limited to sensory processing (or a selective read-out of sensory processing), learning should be confined to the trained stimuli; otherwise learning effects will transfer irrespective of the variations in other stimulus features.

The results indicated a notable specificity to stimulus variations. In the component-training group, we did not find significant transfer for trained and test stimuli that differed in motion directions (Fig. 6E left panel, pre-/post-test, t(7) = 1.886, p = 0.101; Fig. 6F left panel, PI, t(7) = 2.016, p = 0.084). We also found no significant transfer to test stimuli that have smaller size (Fig. 6E left panel, pre-/post-test, t(7) = 1.308, p = 0.232; Fig. 6F left panel, PI, t(7) = 1.376, p = 0.211) or lower contrast (Fig. 6E left panel, pre-/post-test, t(7) = 2.187, p = 0.065; Fig. 6F left panel, PI, t(7) = 1.971, p = 0.089).

Similarly, if component motion directions were switched such that the resulting plaid moves in an orthogonal direction, transfer effects in the plaid-training group were not statistically evident (Fig. 6E right panel, pre-/post-test, t(5) = 1.268, p = 0.261; Fig. 6F right panel, PI, t(5) = 1.645, p = 0.161). We also investigated how changing stimulus speed affects learning transfer. When the grating speed was increased to match the apparent speed of the trained plaid, the transfer effect was not significant (Fig. 6E right panel, pre-/post-test, t(5) = 1.257, p = 0.265; Fig. 6F right panel, PI, t(5) = 1.635, p = 0.163).

Taken together, we find that motion VPL is specific to stimulus direction, speed, size, and contrast. These results demonstrate that our training has strong susceptibilities to variations in basic visual features. Such strong dependencies indicate that a broadly tuned non-sensory learning mechanism unlikely plays an important role in observed learning because it predicts a broad transfer over variations in low-level stimulus features. Note that we cannot completely eliminate the possibility of changes in sensory readout mechanisms since, theoretically, a refined readout mechanism can be sensitive to changes in sensory features[45–47]. Nonetheless, these results suggest the pivotal roles of basic stimulus features in perceptual learning of motion.

Discussion

Elucidating where in the visual processing hierarchy plasticity associated with VPL takes place has been a key question in perceptual learning research over the past decades. Here, we addressed this question in the domain of motion perception. We trained participants to identify motion directions of either component motion (a drifting grating) or pattern motion (a drifting plaid), and assessed transfer of learning to a variety of carefully controlled stimulus conditions. The bidirectional transfer of learning between component and pattern motion provides evidence for learning effects that are associated with the middle-levels of processing where component motions are combined into plaid percepts, and, at the same time, rules out plasticity at the low-levels where complex motions are represented as components. In addition, we also observed specificities to the trained direction, speed, stimulus size, and contrast. These results are in line with the previous findings that VPL is generally vulnerable to the variations in basic feature dimensions and argue against the plasticity in high-level brain areas that represent non-sensory cognitive factors, such as general task statistics and decision rules[15–17].

Our results are of significance for understanding mechanisms underlying motion perception. As one of the key research topics in vision science, dissociable functional roles of the low-level and the middle-level motion system have been well documented[23,48,49]. The seminal paper by Adelson and Movshon[39] documented how moving plaid percepts can arise from component gratings. Subsequent neurophysiological work discovered distinct tuning properties of individual neurons in V1 and MT with preferences toward component and plaid representations, respectively[36]. The degree of bias toward plaid representations in MT appears to depend on whether the animal is anesthetized or not. Namely, MT exhibits stronger bias toward representing plaid motion direction in awake animals[38]. These findings were generalized to humans. Huk and Heeger[50] used functional magnetic resonance imaging (fMRI) and reported robust fMRI adaptation to pattern motion in the human motion-sensitive area hMT + . Thus, the phenomenon of component and pattern motion usually serves as a good benchmark for studying visual hierarchy of motion processing. Note that some scholars also suggested the possibility that neuronal responses in MT might reflect tracking of moving intersections (i.e., motion of luminance blobs caused by adding two gratings) rather than the feedforward encoding of true pattern motion direction[38]. Our stimuli did not allow us to disentangle these two possibilities. Nonetheless, we can rule out general improvement in intersection tracking as plaid training did not transfer when we changed the direction of pattern motion (Fig. 6F, right panel).

In this paper, we took advantage of the known visual motion processing hierarchy and asked how perceptual training might alter behaviorally relevant signals at different motion processing stages. We address this question by showing that training on component or pattern motion bi-directionally transfers to each other if the two stimuli share the same apparent motion direction. These results suggest that, when a plaid motion stimulus is being learned, learning signals might preferentially refine processing associated with pattern-selective units, but not the component sensitive units. While there have been many behavioral studies of motion VPL[7–9,51–53], to our knowledge, no studies employed an experimental design that distinguishes between component- and pattern-dependent representations. For instance, VPL studies typically relied on random-dot-kinematogram stimuli or trained subjects on fine direction discrimination tasks[7,54,55]. Studies that used gratings only tested contrast thresholds for coarse motion direction judgments[52].

Our study also constrains theoretical models of VPL. Two distinct computational frameworks of VPL have emerged so far, where learning either improves the quality of sensory encoding or optimizes high-level readout and decision mechanisms that can in turn promote perceptual sensitivity. Empirical evidence, however, is highly contentious. Early psychophysical studies on motion VPL demonstrated the considerable specificity to the trained direction[7,8], implying plasticity associated with direction-selective units. However, specificities in motion VPL have also been shown to be mediated by other factors, such as task difficulty[51,56], exposure to other directions[53], external noise[52]. This debate in VPL psychophysics is mirrored by a similar debate with respect to the neural substrates of VPL. For example, after training monkeys on a motion direction decision task, Law and Gold[28] found pronounced behaviorally relevant changes in neural responses in area LIP, but minimal changes in neural activity in area MT. This study advocates a mechanism beyond the sensory-representation level, where training results in a more efficient extraction of useful sensory information rather than in an enhancement of sensory representations per se. In contrast, recent fMRI studies found that motion VPL refines the cortical tuning of the hMT+, emphasizing the pivotal role of enhancement at sensory-representation level[57,58]. Notably, the mechanistic role of high-level cognitive influences in sensory processing is still largely unknown. Previous studies have suggested at least two broad types of learning mechanisms, those that are sensory (e.g., selective readout) and those that are non-sensory (e.g., conceptual learning, rule-based learning). While disentangling these higher-level processes is beyond the scope of this paper, the observed specificity to basic stimulus features (Fig. 6) argues against non-sensory cognitive factors.

What are the possible neural underpinnings of the observed empirical findings in the present work? We surmise that several mechanisms may coexist and interact. First, because training on a plaid motion stimulus does not fully transfer to its two components (Fig. 5E), we conclude that a significant part of the relevant plasticity occurs downstream from the low-level motion mechanisms. Given the evidence that MT neurons analyze pattern motion by selectively integrating inputs from a population of V1 neurons[46], one possible mechanism is that learning improves information transmission from the low-level to the middle-level motion processing. Such a mechanism is consistent with findings of a recent study where attention was shown to improve the amount of information transferred from V1 to hMT+[59]. Moreover, learning effects in our study are specific to direction, speed, contrast, and size, indicating critical roles of neuronal tuning to these low-level visual features. For example, stimulus contrast and size have strong influences on neural responses in motion processing[60]. This is also in line with our previous findings showing that motion perception is strongly modulated by stimulus contrast and size[61,62]—behavioral findings that have been linked to mechanisms within area MT[63,64]. Finally, our results are also consistent with stimulus-selective changes in sensory readout mechanisms[45–47]. Note that in all of these possible explanations, critical plasticity involves pattern-selective motion mechanisms albeit in somewhat different contexts.

In summary, our study provides evidence for the training-induced plasticity associated with the intermediate stage of motion processing, and highlights the significance of basic motion-related visual attributes in mediating the transfer of motion VPL.

Data Availability

Data are available from the corresponding author upon reasonable request.

References

1. Sasaki, Y., Nanez, J. E. & Watanabe, T. Advances in visual perceptual learning and plasticity. *Nat. Rev. Neurosci.* **11**, 53–60 (2010).
2. Watanabe, T. & Sasaki, Y. Perceptual learning: toward a comprehensive theory. *Annu. Rev. Psychol.* **66**, 197–221 (2015).
3. Adini, Y., Sagi, D. & Tsodyks, M. Context-enabled learning in the human visual system. *Nature.* **415**, 790–793 (2002).
4. Adini, Y., Wilkonsky, A., Haspel, R., Tsodyks, M. & Sagi, D. Perceptual learning in contrast discrimination: the effect of contrast uncertainty. *J Vis.* **4**, 993–1005 (2004).
5. Yu, C., Klein, S. A. & Levi, D. M. Perceptual learning in contrast discrimination and the (minimal) role of context. *J Vis.* **4**, 169–182 (2004).
6. Fahle, M. Specificity of learning curvature, orientation, and vernier discriminations. *Vision Res.* **37**, 1885–1895 (1997).
7. Ball, K. & Sekuler, R. A specific and enduring improvement in visual motion discrimination. *Science.* **218**, 697–698 (1982).
8. Ball, K. & Sekuler, R. Direction-specific improvement in motion discrimination. *Vision Res.* **27**, 953–965 (1987).
9. Fine, I. & Jacobs, R. A. Perceptual learning for a pattern discrimination task. *Vision Res.* **40**, 3209–3230 (2000).
10. Bi, T., Chen, N., Weng, Q., He, D. & Fang, F. Learning to discriminate face views. *J. Neurophysiol.* **104**, 3305–3311 (2010).
11. Furmanski, C. S. & Engel, S. A. Perceptual learning in object recognition: object specificity and size invariance. *Vision Res.* **40**, 473–484 (2000).
12. Fahle, M. & Edelman, S. Long-term learning in vernier acuity: effects of stimulus orientation, range and of feedback. *Vision Res.* **33**, 397–412 (1993).
13. Xiao, L. Q. *et al.* Complete transfer of perceptual learning across retinal locations enabled by double training. *Curr. Biol.* **18**, 1922–1926 (2008).
14. Wang, R., Zhang, J. Y., Klein, S. A., Levi, D. M. & Yu, C. Task relevancy and demand modulate double-training enabled transfer of perceptual learning. *Vision Res.* **61**, 33–38 (2012).
15. Kahnt, T., Grueschow, M., Speck, O. & Haynes, J. D. Perceptual learning and decision-making in human medial frontal cortex. *Neuron.* **70**, 549–559 (2011).

16. Wang, R. *et al.* Perceptual Learning at a Conceptual Level. *J. Neurosci.* **36**, 2238–2246 (2016).
17. Zhang, J. Y. *et al.* Rule-based learning explains visual perceptual learning and its specificity and transfer. *J. Neurosci.* **30**, 12323–12328 (2010).
18. Maniglia, M. & Seitz, A. R. Towards a whole brain model of Perceptual Learning. *Curr Opin Behav Sci.* **20**, 47–55 (2018).
19. Grill-Spector, K. & Malach, R. The human visual cortex. *Annu. Rev. Neurosci.* **27**, 649–677 (2004).
20. Wandell, B. A. Computational neuroimaging of human visual cortex. *Annu. Rev. Neurosci.* **22**, 145–173 (1999).
21. Park, W. J. & Tadin, D. Motion Perception in *Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience* Vol. 2 *Series* (ed. John Serences) Ch. Sensation, Perception, and Attention, 415–488 (Wiley, 2018).
22. Hubel, D. H. & Wiesel, T. N. Receptive fields and functional architecture of monkey striate cortex. *J. Physiol.* **195**, 215–243 (1968).
23. Born, R. T. & Bradley, D. C. Structure and function of visual area MT. *Annu. Rev. Neurosci.* **28**, 157–189 (2005).
24. Newsome, W. T. & Pare, E. B. A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *J. Neurosci.* **8**, 2201–2211 (1988).
25. Servos, P., Osu, R., Santi, A. & Kawato, M. The neural substrates of biological motion perception: an fMRI study. *Cereb. Cortex.* **12**, 772–782 (2002).
26. Antal, A. *et al.* Direct current stimulation over V5 enhances visuomotor coordination by improving motion perception in humans. *J. Cogn. Neurosci.* **16**, 521–527 (2004).
27. Shadlen, M. N. & Newsome, W. T. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J. Neurophysiol.* **86**, 1916–1936 (2001).
28. Law, C. T. & Gold, J. I. Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nat. Neurosci.* **11**, 505–513 (2008).
29. Latimer, K. W., Yates, J. L., Meister, M. L., Huk, A. C. & Pillow, J. W. Single-trial spike trains in parietal cortex reveal discrete steps during decision-making. *Science.* **349**, 184–187 (2015).
30. Gold, J. I. & Shadlen, M. N. The neural basis of decision making. *Annu. Rev. Neurosci.* **30**, 535–574 (2007).
31. Pilly, P. K., Grossberg, S. & Seitz, A. R. Low-level sensory plasticity during task-irrelevant perceptual learning: evidence from conventional and double training procedures. *Vision Res.* **50**, 424–432 (2010).
32. Watanabe, T. *et al.* Greater plasticity in lower-level than higher-level visual motion processing in a passive perceptual learning task. *Nat. Neurosci.* **5**, 1003–1009 (2002).
33. Glasser, D. M., Tsui, J. M., Pack, C. C. & Tadin, D. Perceptual and neural consequences of rapid motion adaptation. *Proc Natl Acad Sci USA* **108**, E1080–1088 (2011).
34. Ledgeway, T. & Smith, A. T. The duration of the motion aftereffect following adaptation to first-order and second-order motion. *Perception.* **23**, 1211–1219 (1994).
35. Hong, S. W., Tong, F. & Seiffert, A. E. Direction-selective patterns of activity in human visual cortex suggest common neural substrates for different types of motion. *Neuropsychologia.* **50**, 514–521 (2012).
36. Rodman, H. R. & Albright, T. D. Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT). *Exp. Brain Res.* **75**, 53–64 (1989).
37. Movshon, J. A. & Newsome, W. T. Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. *J. Neurosci.* **16**, 7733–7741 (1996).
38. Pack, C. C., Berezovskii, V. K. & Born, R. T. Dynamic properties of neurons in cortical area MT in alert and anaesthetized macaque monkeys. *Nature.* **414**, 905–908 (2001).
39. Adelson, E. H. & Movshon, J. A. Phenomenal coherence of moving visual patterns. *Nature.* **300**, 523–525 (1982).
40. Pelli, D. G. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spat. Vis.* **10**, 437–442 (1997).
41. Foss-Feig, J. H., Tadin, D., Schauder, K. B. & Cascio, C. J. A substantial and unexpected enhancement of motion perception in autism. *J. Neurosci.* **33**, 8243–8249 (2013).
42. Tadin, D., Lappin, J. S. & Blake, R. Fine temporal properties of center-surround interactions in motion revealed by reverse correlation. *J. Neurosci.* **26**, 2614–2622 (2006).
43. Gizzi, M., Newsome, W. & Movshon, J. Directional selectivity of neurons in macaque MT. *Invest. Ophthalmol. Vis. Sci.* **24**, 107 (1983).
44. Movshon, J. A., Albright, T. D., Stoner, G. R., Majaj, N. J. & Smith, M. A. Cortical responses to visual motion in alert and anesthetized monkeys. *Nat. Neurosci.* **6**(3), author reply 3–4 (2003).
45. Dosher, B. A., Jeter, P., Liu, J. & Lu, Z. L. An integrated reweighting theory of perceptual learning. *Proc Natl Acad Sci USA* **110**, 13678–13683 (2013).
46. Rust, N. C., Mante, V., Simoncelli, E. P. & Movshon, J. A. How MT cells analyze the motion of visual patterns. *Nat. Neurosci.* **9**, 1421–1431 (2006).
47. Mollon, J. D. & Danilova, M. V. Three remarks on perceptual learning. *Spat. Vis.* **10**, 51–58 (1996).
48. Simoncelli, E. P. & Heeger, D. J. A model of neuronal responses in visual area MT. *Vision Res.* **38**, 743–761 (1998).
49. Adelson, E. H. & Bergen, J. R. Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am. A.* **2**, 284–299 (1985).
50. Huk, A. C. & Heeger, D. J. Pattern-motion responses in human visual cortex. *Nat. Neurosci.* **5**, 72–75 (2002).
51. Liu, Z. L. Perceptual learning in motion discrimination that generalizes across motion directions. *Proceedings of the National Academy of Sciences of the United States of America.* **96**, 14085–14087 (1999).
52. Lu, Z. L., Chu, W. & Dosher, B. A. Perceptual learning of motion direction discrimination in fovea: separable mechanisms. *Vision Res.* **46**, 2315–2327 (2006).
53. Zhang, J. Y. & Yang, Y. X. Perceptual learning of motion direction discrimination transfers to an opposite direction with TPE training. *Vision Res.* **99**, 93–98 (2014).
54. Liu, Z. & Weinshall, D. Mechanisms of generalization in perceptual learning. *Vision Res.* **40**, 97–109 (2000).
55. Shibata, K. *et al.* Decoding reveals plasticity in V3A as a result of motion perceptual learning. *PLoS One.* **7**, e44003 (2012).
56. Ahissar, M. & Hochstein, S. Task difficulty and the specificity of perceptual learning. *Nature.* **387**, 401–406 (1997).
57. Chen, N. *et al.* Sharpened cortical tuning and enhanced cortico-cortical communication contribute to the long-term neural mechanisms of visual motion perceptual learning. *NeuroImage.* **115**, 17–29 (2015).
58. Chen, N., Cai, P., Zhou, T., Thompson, B. & Fang, F. Perceptual learning modifies the functional specializations of visual cortical areas. *Proc Natl Acad Sci USA* **113**, 5724–5729 (2016).
59. Saproo, S. & Serences, J. T. Attention improves transfer of motion information between V1 and MT. *J. Neurosci.* **34**, 3586–3596 (2014).
60. Pack, C. C., Hunter, J. N. & Born, R. T. Contrast dependence of suppressive influences in cortical area MT of alert macaque. *J. Neurophysiol.* **93**, 1809–1815 (2005).
61. Tadin, D., Lappin, J. S., Gilroy, L. A. & Blake, R. Perceptual consequences of centre-surround antagonism in visual motion processing. *Nature.* **424**, 312–315 (2003).
62. Tadin, D. Suppressive mechanisms in visual motion processing: From perception to intelligence. *Vision Res.* **115**, 58–70 (2015).
63. Schallmo, M. P. *et al.* Suppression and facilitation of human neural responses. *Elife.* **7** (2018).
64. Liu, L. D., Haefner, R. M. & Pack, C. C. A neural basis for the spatial suppression of visual motion perception. *Elife.* **5** (2016).
65. Lappin, J. S., Tadin, D., Nyquist, J. B. & Corn, A. L. Spatial and temporal limits of motion perception across variations in speed, eccentricity, and low vision. *J Vis.* **9**(30), 31–14 (2009).

Acknowledgements

Author Contributions

R.Z. conceived, designed and conducted the study. R.Z. analyzed the data. R.Z. and D.T. wrote the manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-37892-x.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.