Neuroscience

The asymmetric transfers of visual perceptual learning determined by the stability of geometrical invariants

Yan Yang, Yan Zhuo, Zhentao Zuo, Tiangang Zhuo, Lin Chen

State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China • University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing 100049, China • Hefei Comprehensive National Science Center, Institute of Artificial Intelligence, Hefei 230088, China

https://en.wikipedia.org/wiki/Open_access

Copyright information

Reviewed Preprint

Published from the original preprint after peer review and assessment by eLife.

About eLife's process

Reviewed preprint version 1
February 28, 2024 (this version)

Posted to preprint server
January 3, 2024

Sent for peer review
November 28, 2023

Abstract

We could recognize the dynamic world quickly and accurately benefiting from extracting invariance from highly variable scenes, and this process can be continuously optimized through visual perceptual learning. It is widely accepted that more stable invariants are prior to be perceived in the visual system. But how the structural stability of invariants affects the process of perceptual learning remains largely unknown. We designed three geometrical invariants with varying levels of stability for perceptual learning: projective (e.g., collinearity), affine (e.g., parallelism), and Euclidean (e.g., orientation) invariants, following the Klein's Erlangen program. We found that the learning effects of low-stability invariants could transfer to those with higher stability, but not vice versa. To uncover the mechanism of the asymmetric transfers, we used deep neural networks to simulate the learning procedure and further discovered that more stable invariants were learned faster. Additionally, the analysis of the network's weight changes across layers revealed that training on less stable invariants induced more changes in lower layers. These findings suggest that the process of perceptual learning in extracting different invariants is consistent with the Klein hierarchy of geometries and the relative stability of the invariants plays a crucial role in the mode of learning and generalization.

eLife assessment

This **important** study proposes a framework to understand and predict generalization in visual perceptual learning in humans based on form invariants. Using behavioral experiments in humans and by training deep networks, the authors offer evidence that the presence of stable invariants in a task leads to faster learning. However, this interpretation is promising but **incomplete**. It can be strengthened through clearer theoretical justification, additional experiments, and by rejecting alternate explanations.

Introduction

In order to adapt to a continually evolving and changing environment, the organism must identify information that remains consistent and stable in dynamic changes (***Gold and Stocker, 2017***). Through the process of visual perceptual learning (VPL), the visual system acquires an increased ability to extract meaningful and structured information from the environment to guide decisions and actions adaptively as the result of experience (***Adolph and Kretch, 2015***; ***Gibson and Pick, 2000***; ***Gibson, 1969***; ***Gold and Watanabe, 2010***). The "information" mentioned above can be understood as invariance preserved under transformation in perception (***Gibson, 1979***).

It is widely accepted that different kinds of invariant properties hold distinct ecological significance and possess different levels of utility in perception (***Buccella, 2021***). According to Klein's Erlangen Program (***Klein, 1893***), a geometrical property is considered as an invariant preserved over a corresponding shape-changing transformation, the more general a transformation group, the more fundamental and stable the geometrical invariants over this transformation group. Stratify geometrical invariants in ascending order of stability: Euclidean geometry, affine geometry, and projective geometry. A fairly large set of experimental results collected within a variety of paradigms have converged at the conclusion that the relative perceptual salience and priority of different attributes of an object may be systematically related to their structural stability under change in a manner that is similar to the Klein hierarchy of geometries: the more stable the attributes, the higher their perceptual salience and priority (***Chen, 2005***, ***1985***, ***1982***; ***Todd et al., 2014***, ***1998***). However, it is remains unknown that whether the learning of invariants with different stability follows a certain pattern, and this is the major focus of our research. To accomplish this, we must first grasp the characteristics of VPL.

According to Gibson's differentiation view (***Gibson and Gibson, 1955***), perceptual learning is a process of "differentiating previously vague impressions" and whereby perceptual information becomes increasingly specific to the stimuli in the world. They suggested that learning as differentiation is the discovery and selective processing of the information most relevant to a task, and this includes discovery of "higher-order invariants" that govern some classification and filter out irrelevant information.

A hallmark of VPL is its specificity for the basic attributes of the trained stimulus and task in the past for a long time (***Crist et al., 1997***; ***Fiorentini and Berardi, 1981***; ***Hua et al., 2010***). Recent studies have challenged the specificity of learned improvements and demonstrated transfer effects between stimuli (***Liu and Weinshall, 2000***; ***Sowden et al., 2002***; ***Zhang et al., 2010***), location (***Hung and Seitz, 2014***) and substantially different tasks (***McGovern et al., 2012***; ***Szpiro and Carrasco, 2015***). To be of practical utility, the generalization of learning effects should be a research focus, and understanding the determinants of specificity and transfer remains one of the large outstanding questions in field. Ahissar and Hochstein uncovered the relationship between task difficulty and transfer effects (***Ahissar and Hochstein, 1997***), leading to the formulation of the reverse hierarchy theory (RHT) which suggests that VPL is a top-down process that originates from the top of the cortical hierarchy and gradually progresses downstream to recruit the most informative neurons to encode the stimulus (***Ahissar and Hochstein, 2004***). More specifically, the level at which learning occurs is related to the difficulty of the task, easier tasks were learned at higher levels and showed more transfer, than did harder tasks. Extending this work, Jeter and colleagues (***Jeter et al., 2009***) demonstrated that the precision of the trained stimuli, and not task difficulty per se, was the critical factor determining transfer, and that learning in fact transferred to low-precision tasks but not to high-precision tasks.

The research described in the present article concerns the very nature of form perception, trying to explore whether VPL of geometrical invariants with various stability also exhibit hierarchical relationships and what a priori rule defines the mode of learning and generalization. Our research focuses on the VPL of different geometrical properties in the Klein hierarchy of geometries:

projective property (e.g., collinearity), affine property (e.g., parallelism), and Euclidean property (e.g., orientation). We developed two psychophysical experiments assessing how the structural stability of geometrical properties affect the learning effect, meanwhile investigating the transfer effect between different levels of geometrical invariants. To explore the relationship between behavioral learning and plasticity across the visual hierarchy during VPL, we perform an experiment in a deep neural network (DNN) that recapitulates several known VPL phenomena (***Wenliang and Seitz, 2018***). The network reproduced the behavior results and further unveiled the learning speeds and layer changes associated with the stability of invariants. We will then interpret the results based on the Klein hierarchy of geometries and RHT.

Results

Asymmetric transfer effect: The learning effect consistently transferred from low-stability to high-stability invariants

The paradigm of "configural superiority effects" with reaction time measures Forty-four right-handed healthy subjects participated in Experiment 1. We randomly assigned subjects into three groups trained with one invariant discrimination task: the collinearity (colli.) training group (n = 15), the parallelism (para.) training group (n = 15) and the orientation (ori.) training group (n = 14). The paradigm of "configural superiority effects" (***Chen, 2005***) (faster and more accurate detection/discrimination of a composite display than of any of its constituent parts) was adapted to measure the short-term perceptual learning effects of different levels of invariants. As illustrated in **Figure 1A**, subjects performed the odd-quadrant discrimination task in which they were asked to report which quadrant differs from the other three as fast as possible on the premise of accuracy. During the test phases before and after training (Pre-test and Post-test, ***Figure 1B***), subjects performed the three invariant discrimination tasks (colli., para., ori.) at three blocks respectively, the response times (RTs) were recorded to measure the learning effects. To distinguish VPL from programmed learning due to motor learning, a color discrimination task, served as baseline training, were performed before the main experiment (***Figure 1B***).

First of all, to investigate if there was any speed-accuracy trade-off, one-way repeated measures analysis of variance (ANOVA) with task as within-subject factors was conducted on the accuracies collected in the Pre-test phase. A significant main effect of the task was found ($F(2, 129) = 7.977$, $p = 0.0005$, $\eta_p^2 = 0.110$). We performed further post-hoc analysis carrying out paired t-test with FDR correction to examine the differences of accuracies between each pair of tasks. As a result, the accuracies of the collinearity task were significantly higher than that in the parallelism task ($t(43) = 5.443$, $p < 0.0001$, Hedge's $g = 0.917$), and that in the orientation task ($t(43) = 4.351$, $p = 0.0001$, Hedge's $g = 0.574$). There was no difference in accuracy between the parallelism and orientation task ($t(43) = 1.535$, $p = 0.132$, Hedge's $g = 0.214$). Then the same analysis was applied to the RTs in the three tasks prior to training. A significant main effect of the task was found in ANOVA ($F(2, 129) = 59.557$, $p < 0.0001$, $\eta_p^2 = 0.480$). As shown from the post-hoc test, the RTs of collinearity task were significantly faster than that in the parallelism task ($t(43) = 13.374$, $p < 0.0001$, Hedge's $g = 1.945$), and that in the orientation task ($t(43) = 13.333$, $p < 0.0001$, Hedge's $g = 2.295$). The RT of the parallelism task was faster than that of orientation task ($t(43) = 4.179$, $p < 0.0001$, Hedge's $g = 0.416$). Taken together, the collinearity task has the highest accuracy as well as the faster RT among the three tasks, showing no speed-accuracy trade-off in Experiment 1. What's more, the results before training were in line with the prediction from the Klein hierarchy of geometries, which suggest that the more stable invariants possessed higher detectability, resulting in better task performance. The statistical results of the accuracies in Post-test didn't differ from that in Pre-test (***Figure 2— figure Supplement 1***). In the following analysis, only correct trials were used.

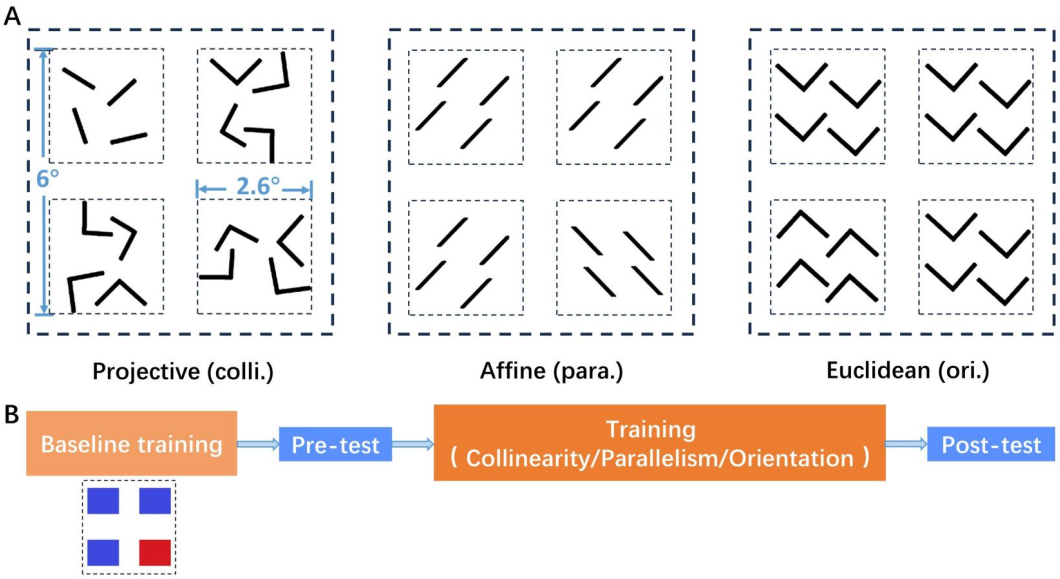


Figure 1.

Examples of stimulus arrays for each task and the procedure in Experiment 1. (A) Stimulus examples of the collinearity discrimination task (left), the parallelism discrimination task (middle), the orientation discrimination task (right). (B) The procedure of Experiment 1.

The second analysis conducted was to assess whether there were learning effects of the trained tasks and transfer effects to the untrained tasks. This was assessed by examining the RTs of Pre-test and Post-test for each geometrical property discrimination task. One-tailed, paired sample t-test was performed to do this. For the collinearity training group, significant learning effect was found ($t(14) = 3.911$, $p = 0.0008$, Cohen's $d = 0.457$), but there was no transfer to the other two untrained task ***(Figure 2A)***. The parallelism training group show significant learning effect ($t(14) = 5.169$, $p < 0.0001$, Cohen's $d = 1.095$), and also show substantial improvement in the collinearity task ($t(14) = 2.753$, $p = 0.008$, Cohen's $d = 0.609$) ***(Figure 2B)***. Moreover, performances of all three task were improved after training on orientation discrimination task: the collinearity task ($t(13) = 4.033$, $p = 0.0007$, Cohen's $d = 0.800$), the parallelism task ($t(13) = 3.482$, $p = 0.002$, Cohen's $d = 0.631$), and the orientation task ($t(13) = 4.693$, $p = 0.0002$, Cohen's $d = 1.048$) ***(Figure 2C)***. This particular pattern of transfer is interesting given the hierarchical relationship of the three different stimulus configurations. For instance, the performance improvement obtained on the parallelism task transferred to the collinearity task which is more stable, whereas not transferred to the orientation task which is less stable. Similarity, learned improvement in orientation discrimination transferred to more stable tasks, with RTs of both collinearity and parallelism tasks showing significant improvements. However, training on collinearity discrimination which is the most stable among the three tasks exhibited task specificity. These findings indicate that perceptual improvements derived from training on a relatively low-stability form invariant can transfer to those invariants with higher stability, but not vice versa.

Finally, we assessed whether learning effect differ in different form invariants. To this end, we computed the " learning index " (LI) ***(Petrov et al., 2011)***, which quantifies learning relative to the baseline performance. ANOVA analysis did not find a significant difference in the LIs among the three tasks ($F(2, 41) = 2.246$, $p = 0.119$, $\eta_p^2 = 0.100$) ***(Figure 2—figure Supplement 2)***.

What needs to be cautious is that, Experiment 1 with RT measures has limitations that make it difficult to truly compare the time required for processing different invariants. Specifically, our interest lies in understanding how learning affects the process of extracting geometrical invariants. However, in order to make a response, participants also need to locate the differing quadrant. The strength of the grouping effect of the shapes among the four quadrants can affect the speed of the localization process ***(Orsten-Hooge et al., 2011)***. Additionally, the strength of the grouping effect may vary under different conditions, leading to differences in reaction times that may reflect differences in the extraction time of geometrical invariants as well as the strength of the group effect among the quadrants.

The paradigm of discrimination with threshold measures To overcome the shortcomings of the RT measures, VPL is indexed by the improvements in thresholds of discrimination tasks after training in Experiment 2. We employed the adaptive staircase procedure QUEST ***(Watson and Pelli, 1983)*** to assess the thresholds. QUEST is a kind of Bayesian adaptive methods which typically produces estimates of psychometric function parameters and converges to a threshold more quickly than conventional staircase procedures. Forty-five healthy subjects participated in Experiment 2, and they were randomly assigned into three groups: the collinearity training group (n = 15), the parallelism training group (n = 15) and the orientation training group (n = 15). The experiment paradigm was adapted from a classical 2-alternative forced choice (2AFC) task, in which subjects were required to judge which of two simultaneously presented stimuli was the "target" ***(Figure 3)***, ***Figure 4*** and ***Figure 3—figure Supplement 1***). The "target" referred to the pair of non-collinear lines for the colli. task, the pair of unparallel lines for the para. task, and the more clockwise line for the ori. task. The trials not involved presentation of a "target" were set as catch trials ***(Figure 3)***. The procedure of Experiment 2 is similar to that of Experiment 1 except for no involvement of the baseline training. For each block, a QUEST staircase was used to adaptively adjust the angle separation (θ) of discrimination task within all trials but the catch trials, and provided an estimate of each subject's 50% correct discrimination threshold.

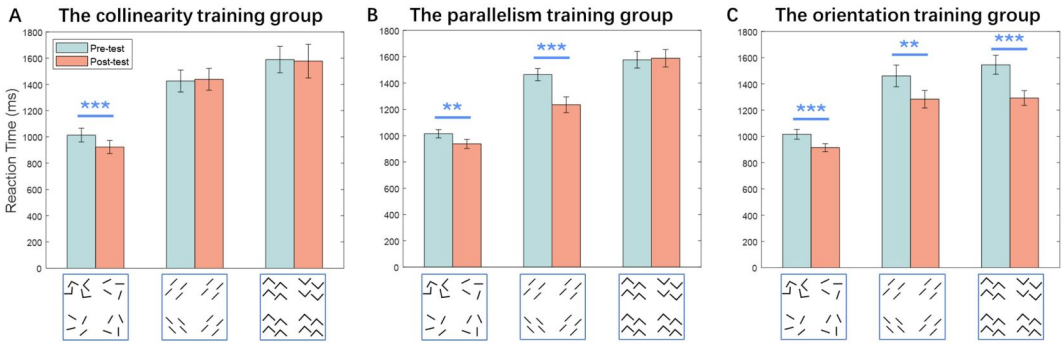


Figure 2.

Results of Experiment 1, RTs of each discrimination tasks measured at Pre-test and Post-test were compared by one-tailed, paired sample t-test. (A) Results from the group trained on the collinearity task (n=15). Performances of the collinearity task were improved after training ($p = 0.0008$). (B) Results from the group trained on the parallelism task (n=15). Performances of the collinearity ($p = 0.008$) and parallelism ($p < 0.0001$) task were improved after training. (C) Results from the group trained on the orientation task (n=14). Performances of the collinearity ($p = 0.0007$), parallelism ($p = 0.002$) and orientation task ($p = 0.0002$) were improved after training. (***$p < 0.001$, **$p < 0.01$, *$p < 0.05$). Error bars denote 1 SEM across subjects.

Figure 2—figure supplement 1. Accuracies for the three discrimination tasks measured at Pre-test and Post-test.

Figure 2—figure supplement 2. The learning indexes of the three geometrical invariants in Experiment 1.

Figure 2—source data 1. RTs and accuracies at Pre-test and Post-test, and learning indexes in the course of training for each participant.

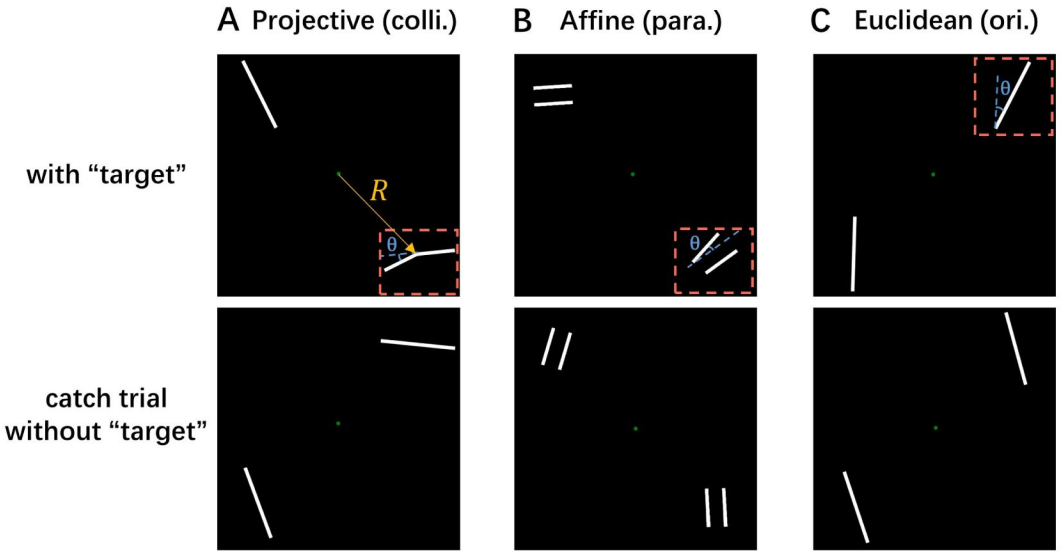


Figure 3.

Examples of the layout of a stimulus frame. The top line demonstrates trials with a “target” (surrounded by orange dashed box), and the bottom line demonstrates the catch trials without “target”. The blue dashed lines represent the “base” orientation for each stimulus, and θ is the angle separation of the discrimination task. (A) Stimulus examples of the collinearity (colli.) task, the upper example shows a “target” (a pair of non-collinear lines) located at the lower right quadrant. (B) Stimulus examples of the parallelism (para.) task, the upper example shows a “target” (a pair of unparallel lines) located at the lower right quadrant. (C) Stimulus examples of the orientation (ori.) task, the upper example shows a “target” (the more clockwise line) located at the upper right quadrant.

Figure 3—figure supplement 1. Examples of stimuli in Experiment 2.

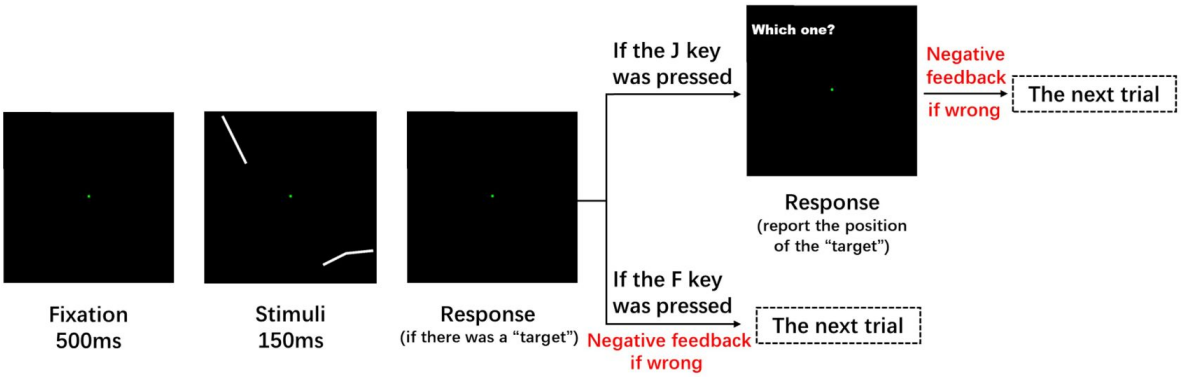


Figure 4.

Examples of stimulus arrays for each task and the procedure in Experiment 1. (A) Stimulus examples of the collinearity discrimination task (left), the parallelism discrimination task (middle), the orientation discrimination task (right). (B) The procedure of Experiment 1

One-way repeated measures ANOVA and post-hoc t-test were conducted on the thresholds collected from all three training groups prior to training. A significant main effect of the task was found in ANOVA analysis ($F(2, 132) = 25.619$, $p < 0.0001$, $\eta_p^2 = 0.280$). As revealed by post-hoc tests, the initial performances of the three tasks were consistent with the relative stability of the invariant they involved, the discrimination threshold of the collinearity task was significantly lower than that of the parallelism task ($t(44) = 3.247$, $p = 0.002$, Hedge's $g = 0.595$), and that of the orientation task ($t(44) = 6.662$, $p < 0.0001$, Hedge's $g = 1.285$). The threshold of the parallelism task was lower than that of the orientation task ($t(44) = 4.570$, $p = 0.0001$, Hedge's $g = 0.916$). Moreover, the accuracies of the three tasks in Pre-test were also submitted to a one-way repeated measures ANOVA and no significant effect was found ($F(2, 132) = 0.046$, $p = 0.955$, $\eta_p^2 = 0.001$), suggesting no difference in difficulty among the three tasks before training.

One-tailed, paired sample t-test was performed to compare the threshold in Pre-test and Posttest. After orientation discrimination training, the orientation discrimination threshold at Post-test was significantly lower than that at Pre-test ($t(14) = 5.527$, $p < 0.0001$, Cohen's $d = 1.516$), and the same applied to the collinearity discrimination task, $t(14) = 2.740$, $p = 0.008$, Cohen's $d = 0.752$, and the parallelism discrimination task, $t(14) = 1.949$, $p = 0.036$, Cohen's $d = 0.654$ ***(Figure 5C)***. After parallelism discrimination training, significant improvements were found in the collinearity discrimination task ($t(14) = 2.775$, $p = 0.007$, Cohen's $d = 1.013$) and the parallelism discrimination task ($t(14) = 3.259$, $p = 0.003$, Cohen's $d = 1.192$) ***(Figure 5B)***. And collinearity discrimination training only produced an improvement on its own performance ($t(14) = 2.128$, $p = 0.026$, Cohen's $d = 0.759$, ***Figure 5A***). In summary, the pattern of generalization is identical to what was found in Experiment 1, where training of low-stability invariants optimized the perception of high-stability invariants but not vice versa. Just the same as in Experiment 1, no differences were found between the LIs of the three tasks in Experiment 2 ($F(2, 42) = 2.875$, $p = 0.068$, $\eta_p^2 = 0.120$, ***Figure 5—figure Supplement 1***).

Previous studies claimed that transfer of VPL is controlled by the difficulty or precision of the training task (***Ahissar and Hochstein, 1997***; ***Jeter et al., 2009***; ***Wenliang and Seitz, 2018***). In this experiment, task difficulty is related to the accuracy, and task precision which is related to the angle separation can be indexed by the threshold. As stated above, prior to training, there was not significant difference among the difficulties (accuracies) of the three tasks, and tasks with higher stability had lower threshold values, resulting in higher precision during training. We showed that the relative stability of invariants determined the transfer effects between tasks even when task difficulty was held constant between tasks, and the particular transfer pattern found in our study (learning from tasks with lower stability and lower precision transferred to the tasks with higher stability and precision) is contrary to the precision-dependent explanation for the generalization of VPL which proposed that training on higher precision can improve performance on lower precision tasks but the reverse is not true (***Jeter et al., 2009***; ***Wenliang and Seitz, 2018***).

Deep neural network simulations of learning and transfer effects

Behavioral results

To gain further insight into the neural basis underlying the asymmetric transfers found in the two psychophysical experiments, in Experiment 3, we repeated Experiment 2 in a DNN for modeling VPL (***Wenliang and Seitz, 2018***). This network, which is derived from the general AlexNet architecture, fulfilled predictions of existing theories (e.g. the RHT) regarding specificity and plasticity and reproduced findings of tuning changes in neurons of the primate visual areas (***Manenti et al., 2023***; ***Wenliang and Seitz, 2018***). Three networks were trained on the three tasks (colli., para., ori.) respectively, repeated in 12 conditions with varying stimulus parameters.

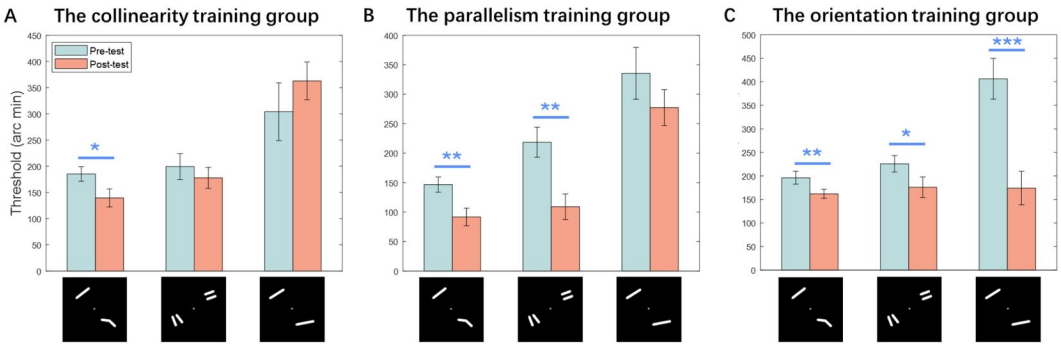


Figure 5.

Results of Experiment 2, Thresholds of each discrimination task measured at Pre-test and Post-test were compared by one-tailed, paired sample t-test. (A) Results from the group trained on the collinearity task (n = 15). Performances of the collinearity task were improved after training (p = 0.026). (B) Results from the group trained on the parallelism task (n = 15). Performances of the collinearity (p = 0.007) and parallelism (p = 0.003) task were improved after training. (C) Results from the group trained on the orientation task (n = 15). Performances of the collinearity (p = 0.008), parallelism (p = 0.036) and orientation task (p < 0.0001) were improved after training. (***p < 0.001, **p < 0.01, *p < 0.05). Error bars denote 1 SEM across subjects.

Figure 5—figure supplement 1. The learning indexes of the three geometrical invariants in Experiment 2.

The performance trajectories of the networks trained with collinearity, parallelism, and orientation discrimination were averaged across the 12 stimulus conditions, and are shown in ***Figure 6A*** respectively. Each network was also tested on the two untrained tasks in the same stimulus condition during the training phase, and the transfer accuracies are presented in ***Figure 6A*** as well. As shown from the final accuracies (the numbers located at the end of each curve), the networks trained on ori. showed greatest transfer effects to the untrained tasks, and the networks trained on colli. showed worst transfer effects.

We then investigate the speeds of learning of the three tasks, we calculated t_{95} (***Wenliang and Seitz, 2018***), the iteration where the fully plastic network reached 95% accuracy, for each task. We found a significant main effect of the training task on t_{95} using the one-way repeated measures ANOVA ($F(2, 33) = 144.636$, $p < 0.0001$, $\eta_p^2 = 0.898$). Further post-hoc analysis with paired t-test showed that the performance of collinearity discrimination started to asymptote earlier than the parallelism ($t(11) = 2.567$, $p = 0.018$, Cohen's $d = 1.012$) and the orientation task ($t(11) = 13.172$, $p < 0.0001$, Cohen's $d = 5.192$). The orientation discrimination had slowest speed of learning, it reached saturation slower than the parallelism task ($t(11) = 11.626$, $p < 0.0001$, Cohen's $d = 4.583$) (***Figure 6B***). That is to say, learning speed was superior for the invariants with higher stability.

Figure 6C shows the final learning and transfer performance on all combinations of training and test task. A linear regression on the final accuracies showed a significant positive main effect of the stability of test task on the performance ($\beta = 0.055$, $t(105) = 2.467$, $p = 0.015$, $R^2 = 0.043$), shown as increasing color gradient from top to bottom. We also found that the stability of training task had a significant negative effect ($\beta = -0.057$, $t(105) = -2.591$, $p = 0.011$, $R^2 = 0.048$), shown as decreasing color gradient from right to left. Overall, consistent with the results in the two psychophysical experiments, these results suggested that transfer is more pronounced from less stable geometrical invariants to more stable invariants than vice versa, shown as higher accuracy on lower-right quadrants compared with top-left quadrants.

Distribution of learning across layers

To demonstrate the distribution of learning over different levels of hierarchy, we next examined the time course of learning across the layers, the weight changes for each layer were shown in ***Figure 7A***. Overall, training on lower-stability geometrical invariants produced greater overall changes. Due to this mismatch of weight initialization, we focus on layers 1–5 with weights initialized from the pre-trained AlexNet.

To characterize learning across layers, we studied when and how much each layer changed during training. First, to quantify when significant learning happened in each layer, we estimated the iteration at which the gradient of a trajectory reached its peak (peak speed iteration, PSI; shown in **Figure 7B**). As the result of a linear regression analysis, in layers 1–5, we observed significant negative main effects of the stability of training task ($\beta = -27.315$, $t(176) = -6.623$, $p < 0.0001$, $R^2 = 0.072$), layer number ($\beta = -17.327$, $t(176) = -6.450$, $p < 0.0001$, $R^2 = 0.065$) and a positive interaction of the two on PSI ($\beta = 6.579$, $t(176) = 5.290$, $p < 0.0001$, $R^2 = 0.115$), suggesting that layer change started to asymptote later for lower layers and less stable invariants. For individual tasks, a linear regression analysis showed a significant negative effect of layer number on PSI only in the least stable task, that is the orientation discrimination task ($\beta = -12.833$, $t(58) = -4.470$, $p < 0.0001$, $R^2 = 0.243$). Therefore, for the discrimination of the least stable invariants, the order of change across layers is consistent with the RHT prediction that higher visual areas change before lower ones (***Ahissar and Hochstein, 2004***, ***1997***).

The final layer changes at the end of training for the networks trained on the three tasks are shown in **Figure 7C** respectively. A linear regression analysis on the final changes in layers 1–5 revealed significant negative main effects of the stability of training task ($\beta = -0.037$, $t(176) = -16.409$, $p < 0.0001$, $R^2 = 0.689$), layer number ($\beta = -0.011$, $t(176) = -7.655$, $p < 0.001$, $R^2 = 0.0001$) and a positive interaction of the two ($\beta = 0.005$, $t(176) = 7.329$, $p < 0.0001$, $R^2 = 0.075$). However, for

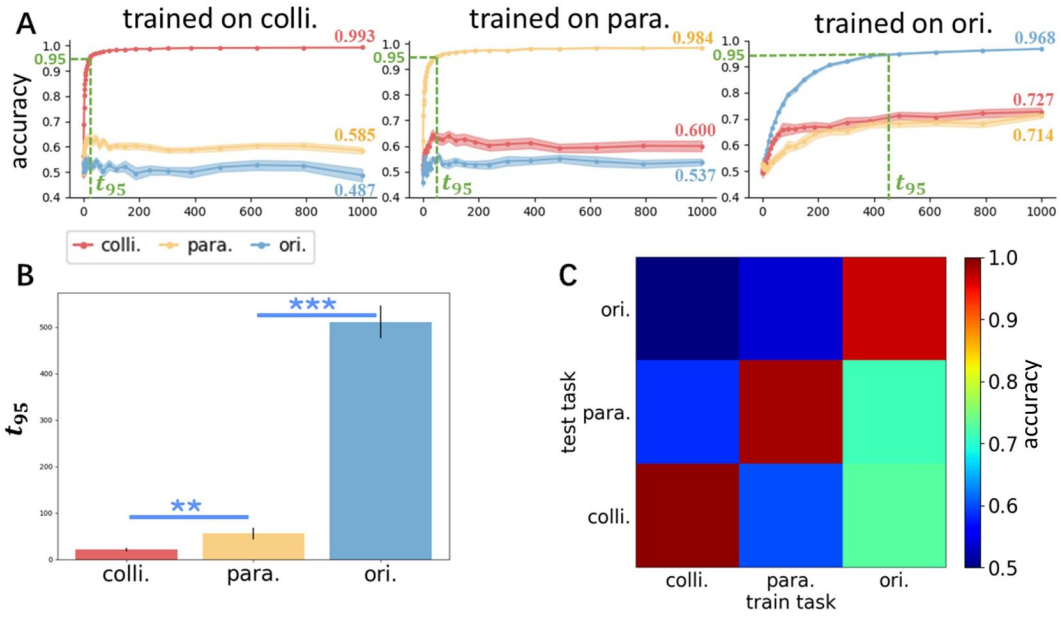


Figure 6.

Performance of the model when trained under different discrimination tasks. (A) Accuracy trajectories against training iterations from the models trained on collinearity (left), parallelism (middle), and orientation task (right), with the error bar representing 1 SEM. t_{95} is the iteration where the fully plastic network reached 95% accuracy, depicted by green dashed lines. The numbers located at the end of each curve are the final accuracies of the last iteration. (B) The learning speed which was indexed by t_{95} of the three tasks. The learning speed of the collinearity task was faster than the parallelism ($p = 0.018$) and orientation task ($p < 0.0001$). The learning speed of the parallelism task was faster than the orientation task ($p < 0.0001$). Statistical significance was calculated by paired t-test with FDR correction. (***$p < 0.001$, **$p < 0.01$, *$p < 0.05$). Error bars denote 1 SEM across subjects. (C) Final mean accuracies when the network was trained and tested on all combinations of tasks

Figure 6—figure supplement 1. Model structure and stimulus examples in Experiment 3.

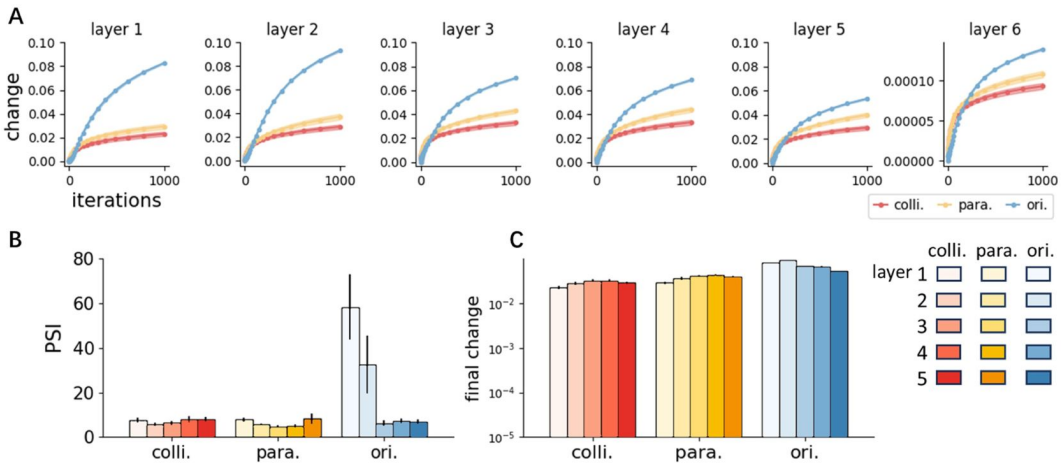


Figure 7.

Layer change under different training tasks. (A) Layer change trajectories during learning. (B) Iteration at which the rate of change peaked (PSI) in layers 1-5. (C) Final layer change in layers 1-5. The error bar representing 1 SEM.

individual tasks, a significant negative linear effect of layer number was only found in the orientation discrimination task (β = -0.008, t(58) = -12.763, p < 0.0001, R^2 = 0.733). On the contrary, significant positive effect of layer number was found in the parallelism (β = 0.003, t(58) = 3.701, p = 0.0005, R^2 = 0.177) and collinearity task (β = 0.002, t(58) = 2.538, p = 0.014, R^2 = 0.084). We then calculated the centroid, which is equal to the weighted average of the layer numbers using the corresponding mean final change as the weight. The centroids for the networks trained on colli., para. and ori. are 3.11, 3.14 and 2.77, respectively. These centroids together with the results from linear regression collectively indicate that the least stable task induces more change lower in the hierarchy whereas the two more stable tasks induce change higher in the hierarchy.

Wenliang and Seitz proposed that high-precision training transfers more broadly to untrained and coarse discriminations than low-precision training (***Wenliang and Seitz, 2018***). However, in each stimulus condition of Experiment 3, the angle separations (precisions) in the collinearity and parallelism task were always the same, and were half of the angle separations in the orientation task. So, the pattern of transfer and layer change cannot be explained based on the relative precision of the training and test tasks that was suggested by Wenliang and Seitz. Rather, the relative stability of invariants involved in the tasks provides consistent and reasonable explanation for the asymmetric transfers found in our study.

Discussion

We find a consistent pattern underlying learning and generalization across the three geometrical invariants in the Klein hierarchy: the less stable invariants have slower learning speeds (revealed by the DNN experiment), and transfer is more pronounced from less stable invariants to more stable invariants than vice versa.

We explain results on the basis of the Klein hierarchy of geometries. First of all, because more stable invariants possess higher perceptual salience and greater ecological significance (***Meng et al., 2019***; ***Todd et al., 2014***), they are not only perceived earlier but also learned earlier, leading to faster learning speeds. We further propose the hypothesis that learning high-stability invariants must precede the learning of low-stability invariants, so the learners have to go through learning of high-stability invariant discrimination to get to the point where learning of low-stability invariant discrimination is accessible. What's more, according to the inclusive relationship between the invariants in terms of their structural stability (***Klein, 1941***), the change of a high-stability invariant includes changes of invariants which are less stable than it, thus the improved ability to discriminate low-stability invariants obtained from training could assist in the discrimination of more stable ones. On the other hand, form invariants with high stability are represented in a holistic manner during learning, their individual components are hard to encoded solely. When performing a form discrimination task, the learners extract the most stable invariants without necessarily extracting their embedded invariants which are less stable. In other words, the less stable invariants are ignored and suppressed when they are embedded into more stable configurations, leading to limited generalization to the less stable invariants via training on more stable invariants. Taken together, the generalization of learning is more substantial from low-stability to high-stability invariants.

According to the distribution of learning across layers found in the DNN experiment, for the orientation task with lowest stability, the order of change across layers is consistent with the RHT prediction that higher visual areas change before earlier ones (***Ahissar and Hochstein, 2004***; ***Hochstein and Ahissar, 2002***), meanwhile, low-stability invariant training induces more change lower in the hierarchy and high-stability training induces more change higher in the hierarchy. Based on these findings, the asymmetric transfers and the distribution of learning across layers can be both explained from the perspective of RHT accompanied by the relative stability of different invariants: VPL is a process that occurs from high-to-low-level visual areas, and from

high-to-low-stability invariants. The VPL of high-stability invariants occurs earlier and relies more on higher-level cortical areas, while the learning of low-stability invariants leads to a greater reliance on lower-level cortical areas that have higher resolution for finer and more difficult discriminations but not involved during the learning of high-stability invariants. On account of the feedforward anatomical hierarchy of the visual system (***Markov et al., 2013***), the modifications of lower areas caused by training on low-stability invariants will also affect higher-level visual areas, thus influencing the discrimination of high-stability invariants and resulting in transfer effects. Taken together, discriminating invariants with higher-stability can benefits from VPL of discriminating invariants with lower-stability, but not vice versa.

Returning to Gibson's theory of perceptual learning, we make a reasonable inference that the Klein hierarchy of geometrics belong to what Gibson referred to as "structure" and "invariant" (***Gibson, 1970***; ***Szokolszky et al., 2019***). More importantly, the invariants with higher structural stability are equivalent to "higher-order invariants" mentioned by Gibson (***Gibson, 1971***), and they are extracted earlier in both processes of perception and perceptual learning. Learning is a process of differentiation, starting with the extraction of global, more stable invariants and gradually involving local, less stable invariants. During this process, the perceptual system becomes more differentiated, more specific, and better at distinguishing local details. Moreover, such a perceptual system can also utilize local information to improve the performance of discriminating high-stability invariants, leading to the asymmetric transfers observed in our research.

We do not deny that several attributes of task including difficulty and precision could play some roles in generalization and the locus of learning, as mentioned in the Introduction. Broadly speaking, there are consistencies between these attributes and the Klein hierarchy of geometries. For example, the more stable form invariant should be more coarse and easier to discriminate than the less stable one in the same presentation condition. However, our study found consistent learning and transfer pattern with the Klein hierarchy of geometries even after controlling for difficulty and precision in Experiment 2 and 3. So, it seems that the relative stability of form invariants is a more essential and determinant factor underlying the transfer of learning effects in our research.

Due to the employment of short-term perceptual learning paradigms in both psychophysical experiments in our study, it has been challenging to accurately track the temporal processes of learning (***Yang et al., 2022***).It is also possible that the lack of observed differences in learning effects between tasks could be due to insufficient learning in some tasks, considering the differences in learning speed among the three tasks, as well as the possibility that different training tasks may involve different short-term and long-term learning processes (***Aberg et al., 2009***; ***Mascetti et al., 2013***).

In future research, it may be beneficial to consider employing long-term perceptual learning to investigate the learning rates of discriminating geometrical invariants with variable stability. Additionally, various brain imaging and neurophysiological techniques should be utilized to study the perception and learning of these form invariants, in order to explore their underlying neural mechanisms. This may contribute to our understanding of object recognition, conscious perception and perceptual development.

Methods and Materials

Participants and apparatus

A total of 89 right-handed healthy subjects participated in this study: 44 in Experiment 1 (24 female, mean age 23.70 ± 3.46 years), and 45 in Experiment 2 (26 female, mean age 23.02 ± 3.40 years). All subjects were naïve to the experiment with normal or corrected-to-normal vision. Subjects provided written informed consent, and were paid to compensate for their time. Sample size was determined based on power calculations following a pilot study showing significant learning effect of the collinearity task for effect size of Cohen's d = 0.728 at 80% power. In each experiment, subjects were randomly assigned to one of three training groups (colli., para., and ori. trainging group). The study was approved by the ethics committee of the Institute of Biophysics at the Chinese Academy of Sciences, Beijing.

Stimuli were displayed on a 24-inch computer monitor (AOC VG248) with a resolution of 1920 × 1080 pixels and a refresh rate of 100 Hz. The experiment was programmed and run in MATLAB (The Mathwork corp, Orien, USA) with the Psychotoolbox-3 extensions (***Brainard, 1997***; ***Pelli, 1997***). Subjects were stabilized using a chin and head rest with visual distance of 70 cm in a dim ambient light room.

Stimuli and tasks for psychophysics experiments

In Experiment 1, we applied the paradigm of "configural superiority effects" (CSEs) to measure the learning effects. CSEs refer to the findings that configural relations between simple components rather than the components themselves may play a basic role in visual processing and were originally revealed by an odd quadrant task, as illustrated in **Figure 1A**. This paradigm was also adapted to measure the relative salience of different levels of invariants (***Chen, 2005***). There were three discrimination tasks: discriminations based on a difference in collinearity (colli., a kind of projective property, shown in **Figure 1A**, left), a difference in parallelism (para., a kind of affine property, shown in **Figure 1A**, middle), a difference in orientation of angles (ori., a kind of Euclidean property, shown in **Figure 1A**, right). The stimuli are composed of white line segments with luminance of 38.83 cd/m^2 presented on black background with luminance of 0.13 cd/m^2. The stimulus array is consisted of four quadrants (visual angle 2.6° × 2.6° for each quadrant) and presented in the center region (visual angle, 6° × 6°). The subjects need to identify which quadrant is different from the others. Either of the two states of an invariant may serve as a target. For example, in the Euclidean invariant condition (**Figure 1A**, right), both upward and downward arrows could be the target. A green central fixation point (RGB [0,130,0], 0.15°) was presented throughout the entire block. Subjects were instructed to perform the actual task while maintaining central fixation. Each trial began immediately after the Space key was pressed. The stimulus array was presented until the subject indicated the location of the odd quadrant ("target") via a manual button press as fast as possible on the premise of accuracy. The response time (RT) in each trial was calculated from the onset of the stimulus array. A negative feedback tone was given if the response was wrong.

The sample stimuli and the layout of a stimulus frame in Experiment 2 are illustrated in **Figure 3** and **Figure 3—figure Supplement 1**, respectively. All stimuli are white (38.83 cd/m^2) presented on black background (0.13 cd/m^2). Each stimulus is composed of a group of line(s): a pair of lines which are collinear or non-collinear in colli. task, a pair of lines which are parallel or unparallel in para. task, and a single long line in ori. Task (**Figure 3—figure Supplement 1**). The stimuli for the three tasks were made up of exactly the same line-segments. Thus, line-segments as well as all local features based on these line-segments, such as luminous flux, and spatial frequency components, were well controlled. The length of line (l) in colli. and para. task is 80 arc min, which is half of the length of line ori. task. The width of line is 2 arc min. The distance (d) between the pair of lines in the parallelism task is 40 arc min. The "base" orientation (the dashed line in **Figure 3** and **Figure 3—figure Supplement 1**) was randomly selected from 0-180°. For

colli. and para., the "base" orientation for each stimulus on a stimulus frame was selected independently. Individual stimulus could occur in one of four quadrants, approximately 250 arc min of visual angle from fixation (*R*); two stimuli were presented at two diagonally opposite quadrants on each trial (**Figure 3**).

Schematic description of a trial in Experiment 2 is shown in **Figure 4**, to make the first-order choice, subjects were instructed to press the J key if they thought there was a "target" as defined by each task, and they should press the F key if the contrary is the case (that is, there was no "target" in this trial). A second-order choice needed to be made if subjects have pressed the J key in the first-order choice: they should select which one was the "target" and report its position by pressing the corresponding key. The response of a trial was regarded as correct only if both choices were correct. This type of experimental design with two-stage choice options together with relatively higher proportion of catch trials (one third of all trials) wound help reduce false alarms and response bias.

Procedure for psychophysics experiments

The overall procedure of Experiment 1 is show in **Figure 1B**. The main VPL procedure consisted of three phases: pre-training test (Pre-test), discrimination training (Training), and post-training test (Post-test). During the test phases, the three form invariant discrimination tasks were performed counterbalance across subjects at three blocks, respectively. During the training phase, all subjects were required to finish 10 blocks of the training task which is determined by their group. Each block contained 40 trials. Before all of the phases, subjects practiced 5 trials per task to make sure that they fully understood the tasks.

The procedure of Experiment 2 is similar to that of Experiment 1 except for no involvement of the baseline training. During each block, subject's threshold was measured for each of the three tasks using a QUEST staircase of 40 trials augmented with 20 catch trials, leading to overall 60 trials per block. During the test phases, the tests for the three tasks were counterbalanced across subjects. The training phase contained 8 blocks of the training task corresponding with the group of subjects. Each subject practiced one block per task with large angle difference to make sure that they fully understood the tasks.

Deep neural network simulations

The deep learning model used in this paper was adopted from Wenliang and Seitz (***Wenliang and Seitz, 2018***). The model was implemented in PyTorch (version 2.0.0) and consists of two parallel streams, each encompassing the first five convolutional layers of AlexNet (***Krizhevsky et al., 2017***) plus one fully connected layer which gives out a single scalar value. The network performed a two-interval two-alternative forced choice (2I-2AFC) task. One stream accepted one standard stimulus and the other stream accepted one comparison stimulus. The comparison stimulus was then compared to the standard stimulus. After the fully connected layers, the outputs of the two parallel streams – two scalar values – were entered to a sigmoid layer to give out one binary value indicating which stimulus was noncolinear, unparallel, or more clockwise, equivalent to the "target" in Experiment 2. Weights at each layer are shared between the two streams so that the representations of the two images are generated by the same parameters.

For the stimulus images, 12 equally spaced "base" orientations were chosen (0-165°, in steps of 15°), and the "base" orientation of the standard and comparison was chosen independently except for the orientation discrimination task. We trained the network on all combinations of the 3 parameters: (1) angle separation between standard and comparison —— ① 5° for colli. & para. and 10° for ori., ② 10° for colli. & para. and 20° for ori., ③ 20° for colli. & para. and 40° for ori.; (2) distance between the pair of lines for para. —— ① 30 pixels, ② 40 pixels; (3) the location of the gap on line for colli. —— ① the midpoint; ② the front one-third. So there were overall 12 (3 × 2 × 2) stimulus conditions. It should be noted that the angle separations in the orientation task were

always twice the angle separations in the other two tasks in each condition, this proportion was based on the initial threshold observed in the behavioral experiment. Here are the other stimulus parameters: length of line in para. (100 pixels); length of line in colli. and ori. (200 pixels); width of line (3 pixels); radius of gap for colli. (5 pixels). Each stimulus was centered on an 8-bit 227 × 227 pixel image with black background.

Three networks were trained on the three tasks (colli., para., ori.) respectively, repeated in 12 stimulus conditions. Each network was trained for 1000 iterations of 60-image batches with batch size of 20 pairs, and meanwhile was tested on the other two untrained tasks. Learning and transfer performances were measured at 30 approximately logarithmically spaced iterations from 1 to 1000. We used the same feature maps and kernel size as the original paper. Network weights were initialized such that the last fully connected layer was initialized by zero and weights in the five convolutional layers were copied from an AlexNet trained on object recognition (downloaded from *http://dl.caffe.berkeleyvision.org/* bvlc_reference_caffenet.caffemodel) to mimic a (pretrained) adult brain. Training parameters were set as follow: learning rate = 0.0001, momentum = 0.9. The cross-entropy loss function was used as an objective function and optimized via stochastic gradient descent.

Data analysis

All data analyses were carried out in Matlab (The Mathwork corp, Orien, USA) and Python (Python Software Foundation, v3.9.13). No data were excluded.

Human behavioral data were analyzed using analysis of variance (ANOVA), post-hoc test, and paired sample t-test. For ANOVAs and post-hoc tests, we computed η_p^2 and Hedge's g as effect sizes. For paired sample t-tests, we computed Cohen's d as effect size. To quantify learning effect, we computed the Learning Index (LI) (***Petrov et al., 2011***) as follows:

$$LI = \frac{Performance_{Pre-test} - Performance_{Post-test}}{Performance_{Pre-test}} \quad (1)$$

where *perf ormance* refers to the RT or threshold obtained at the test phases.

In DNN experiment, Ordinary Least Squares (OLS) method of linear regression was implemented to analyze the transfer effects between different tasks. The following equation describes the model's specification:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon \quad (2)$$

where y represents the dependent variable (final accuracy), x_1 and x_2 represent the two features (training and test task) respectively. β_0 is the intercept, β_1, β_2 are coefficients of the linear model, and ϵ is the error term (unexplained by the model).

To estimate the learning effects in layers of DNN, the differences in weights before and after training at each layer were measured. Specifically, for a particular layer with N total connections to its lower layer, we denote the original N-dimensional weight vector trained on object classification as w (N and w are specified in AlexNet), the change in this vector after perceptual learning as δw, and define the layer change as follows:

$$d_{rel} = \frac{\sum_i^N |\delta w_i|}{\sum_i^N |w_i|} \quad (3)$$

where i indexes each element in the weight vector. Under this measure, scaling the weight vector by a constant gives the same change regardless of dimensionality, reducing the effect of unequal weight dimensionalities on the magnitude of weight change. For the weights in the final readout layer that were initialized with zeros, the denominator in Equation 1 was set to N, effectively measuring the average change per connection in this layer. Due to the convolutional nature of the layers 1–5, d_{rel} is equal to the change in filters that are shared across location in those layers. When comparing weight change across layers, we focus on the first five layers unless otherwise stated.

OLS method of linear regression with interaction terms was implemented to analyze the learning effects across layers. The following equation describes the model's specification:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \epsilon \quad (4)$$

where y represents the dependent variable (PSI or final layer change), x_1 and x_2 represent the two features (training task and layer number) respectively. β_0 is the intercept, β_1, β_2 and β_3 are coefficients of the linear model, and ϵ is the error term (unexplained by the model).

Code availability

The deep neural network is available from the original authors on GitHub (*https://github.com/kevin-w-li/DNN_for_VPL*).

Acknowledgements

The study was funded by grants from Ministry of Science and Technology of China grant, the University Synergy Innovation Program of Anhui Province, and the Chinese Academy of Sciences grants.

Additional information

Funding

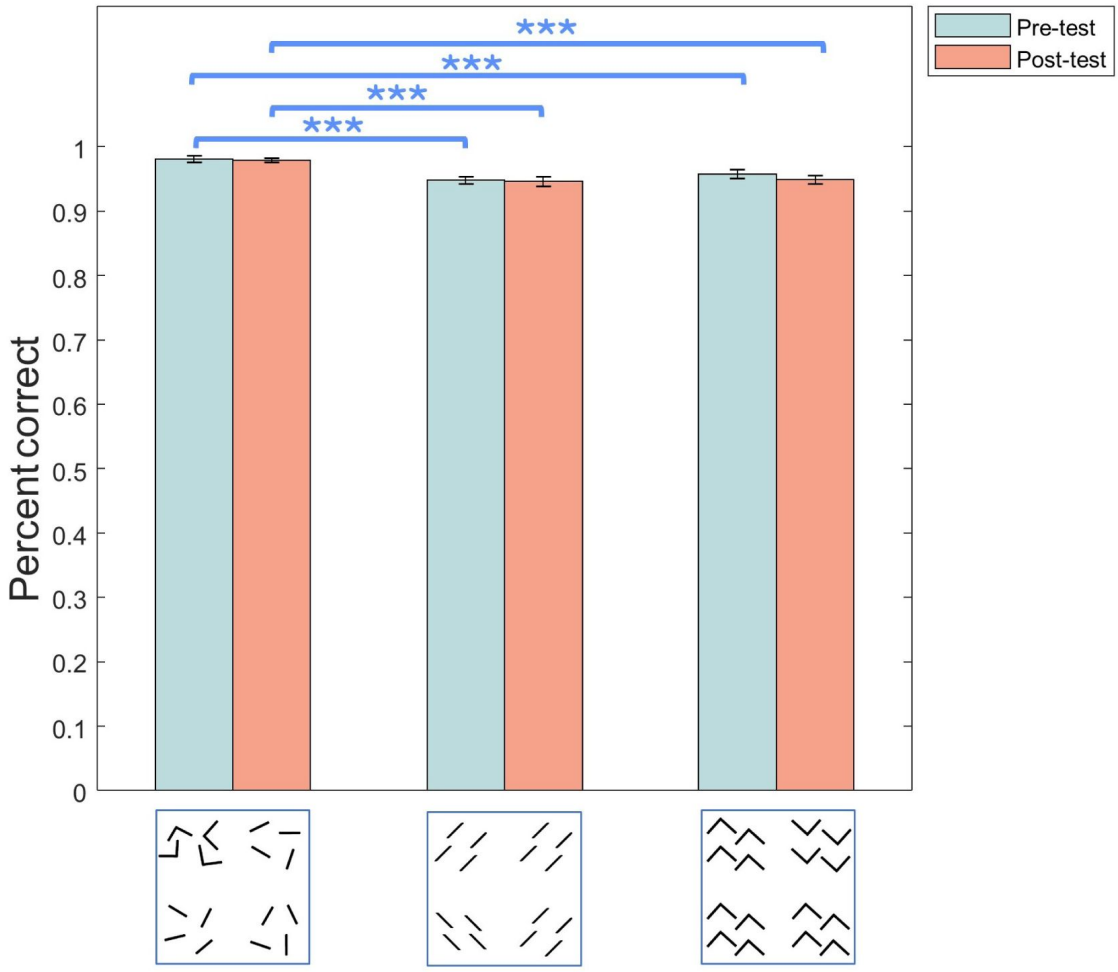
Ministry of Science and Technology of China grant (2020AAA0105601), Tiangang Zhou; Ministry of Science and Technology of China grant (2019YFA0707103 and 2022ZD0209500), Zhentao Zuo; University Synergy Innovation Program of Anhui Province (GXXT-2021-002, GXXT-2022-09), Zhentao Zuo; Chinese Academy of Sciences grants (ZDBS-LY-SM028, 2021091 and YSBR-068), Zhentao Zuo. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Yan Yang, Conceptualization, Data curation, Software, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing – original draft, Project administration, Writing – review & editing; Zhentao Zuo, Tiangang Zhou, Conceptualization, Resources, Supervision, Funding acquisition, Methodology, Project administration, Writing – review & editing; Yan Zhuo, Conceptualization, Resources, Supervision, Funding acquisition; Lin Chen, Conceptualization, Resources, Funding acquisition, Methodology.

Ethics

Human subjects: Informed consent, and consent to publish was obtained from each observer before testing. The study was approved by the ethics committee of the Institute of Biophysics at the Chinese Academy of Sciences, Beijin (reference number:2017-IRB-004).



625

Figure 2—figure supplement 1.

Accuracies for the three discrimination task measured at Pre-test and Post-test. Accuracy was defined as the average percentage correct per block. At Pre-test, the accuracies of the collinearity task were significantly higher than that in the parallelism ($p < 0.0001$) and orientation task ($p = 0.0001$). At Post-test, the accuracies of the collinearity task were still significantly higher than the other two tasks ($p < 0.0001$ for the parallelism task, $p = 0.0001$ for the orientation task). Statistical significance was calculated by paired t-test with FDR correction. (***$p < 0.001$, **$p < 0.01$, *$p < 0.05$). Error bars denote 1 SEM across subjects.

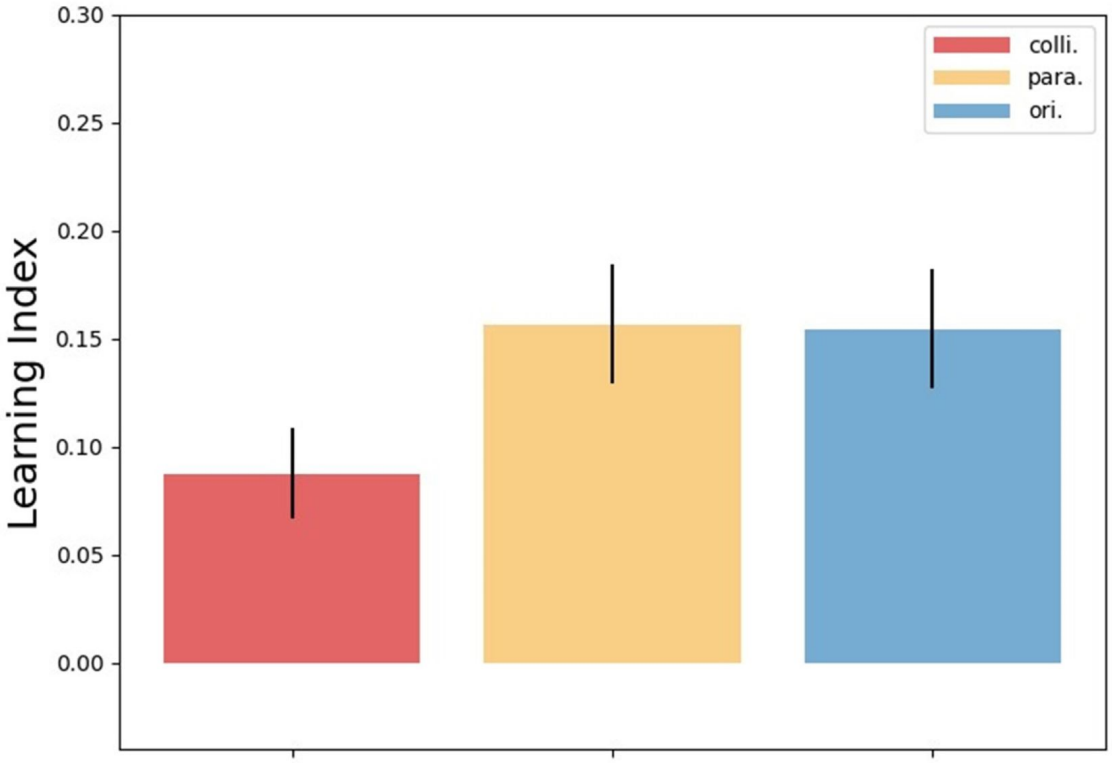


Figure 2—figure supplement 2.

The learning indexes of the three geometrical invariants in Experiment 1. Error bars denote 1 SEM across subjects.

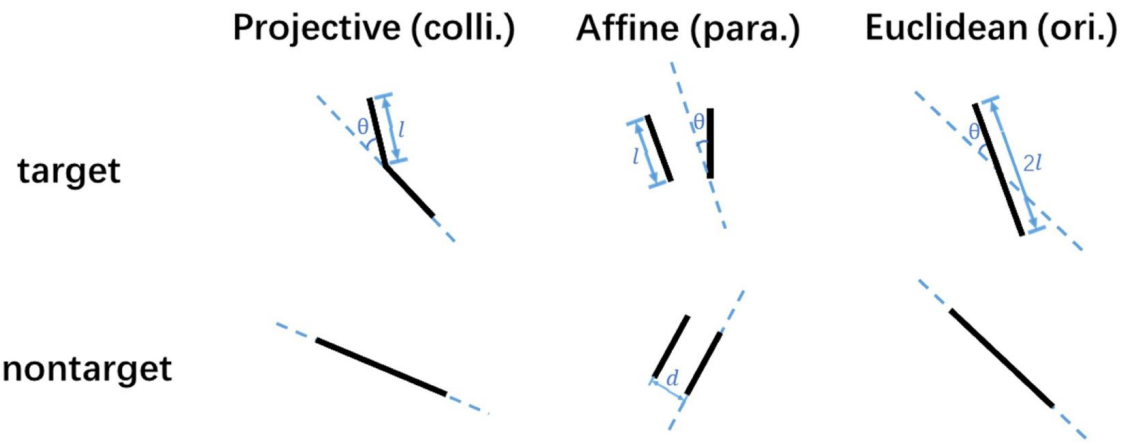


Figure 3—figure supplement 1.

Examples of stimuli in Experiment 2. Sample stimuli in the collinearity (left), parallelism (middle) and orientation (right) discrimination task. The blue dashed lines represent the "base" orientation for each stimulus, and θ is the angle separation of the discrimination task.

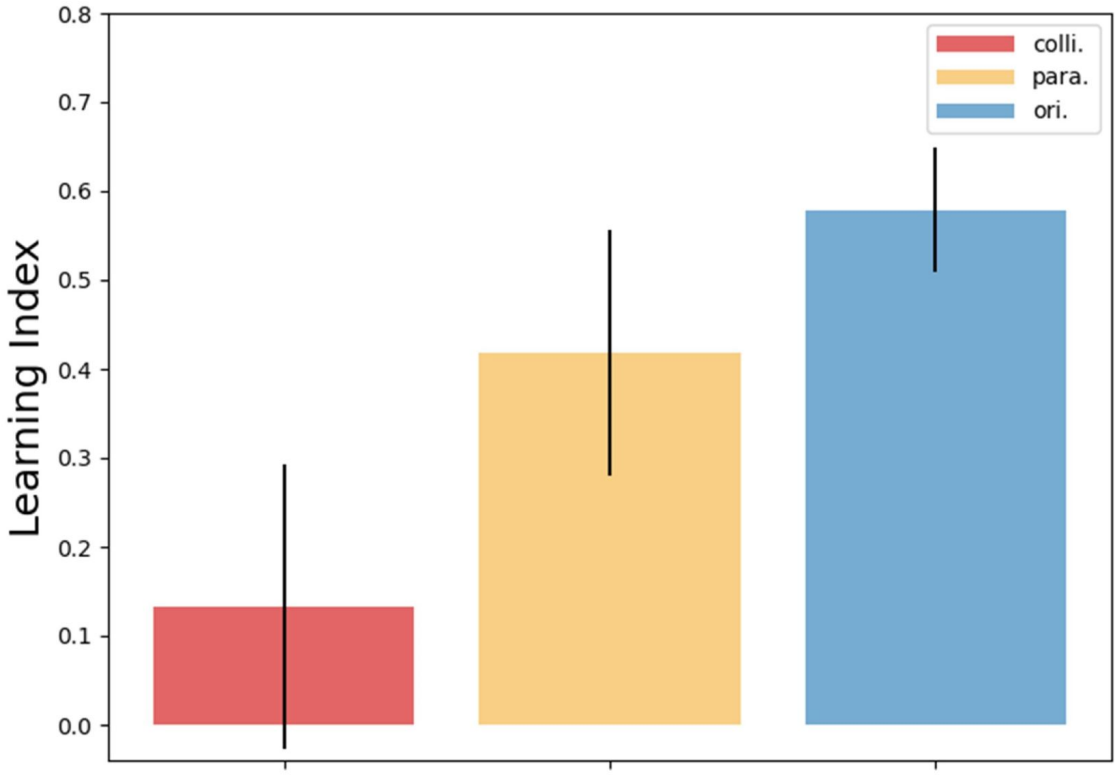


Figure 5—figure supplement 1.

The learning indexes of the three geometrical invariants in Experiment 2. Error bars denote 1 SEM across subjects.

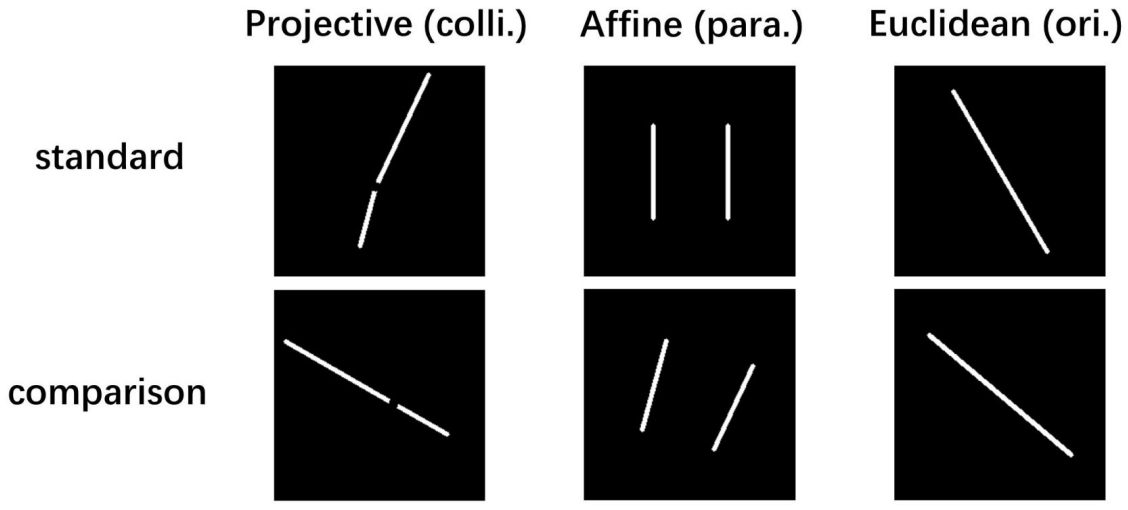


Figure 6—figure supplement 1.

Stimulus examples in Experiment 3. Examples of the pairs of stimulus images for the three discrimination tasks in Experiment 3. The examples here are selected from the stimulus condition with the following parameters: angle separation (10° for colli. & para. and 20° for ori.), distance for para. (40 pixels), location of gap for colli. (the front one-third).

References

Aberg KC, Tartaglia EM, Herzog MH (2009) **Perceptual learning with Chevrons requires a minimal number of trials, transfers to untrained directions, but does not require sleep** *Vision Research* **49**:2087–2094 https://doi.org/10.1016/j.visres.2009.05.020

Adolph KE, Kretch KS (2015) **Gibson's Theory of Perceptual Learning** *International Encyclopedia of the Social & Behavioral Sciences* **12**:127–134 https://doi.org/10.1016/B978-0-08-097086-8.23096-1

Ahissar M, Hochstein S (1997) **Task difficulty and the specificity of perceptual learning** *Nature* **387**:401–406 https://doi.org/10.1038/387401a0

Ahissar M, Hochstein S (2004) **The reverse hierarchy theory of visual perceptual learning** *Trends in Cognitive Sciences* **8**:457–464 https://doi.org/10.1016/j.tics.2004.08.011

Brainard DH. (1997) **The Psychophysics Toolbox** *Spatial Vision* **10**:433–436 https://doi.org/10.1163/156856897X00357

Buccella A (2021) **The problem of perceptual invariance** *Synthese* **199**:13883–13905 https://doi.org/10.1007/s11229-021-03402-2

Chen L (1982) **Topological Structure in Visual Perception** *Science* **218**:699–700 https://doi.org/10.1126/science.7134969

Chen L (1985) **Topological structure in the perception of apparent motion** *Perception* **14**:197–208 https://doi.org/10.1068/p140197

Chen L (2005) **The topological approach to perceptual organization** *Visual Cognition* **12**:553–637 https://doi.org/10.1080/13506280444000256

Crist RE, Kapadia MK, Westheimer G, Gilbert CD (1997) **Perceptual Learning of Spatial Localization: Specificity for Orientation, Position, and Context** *Journal of Neurophysiology* **78**:2889–2894 https://doi.org/10.1152/jn.1997.78.6.2889

Fiorentini A, Berardi N (1981) **Learning in grating waveform discrimination: Specificity for orientation and spatial frequency** *Vision Research* **21**:1149–1158 https://doi.org/10.1016/0042-6989(81)90017-1

Gibson EJ. (1969) **Principles of perceptual learning and development**

Gibson EJ (1970) **The development of perception as an adaptive process** *American scientist* **58**:98–107

Gibson EJ, Pick A (2000) **An Ecological Approach to Perceptual Learning and Development**

Gibson EJ (1971) **Perceptual learning and the theory of word perception** *Cognitive Psychology* **2**:351–368 https://doi.org/10.1016/0010-0285(71)90020-X

Gibson JJ, Gibson EJ (1955) **Perceptual learning: Differentiation or enrichment?** *Psychological Review* **62**:32–41 https://doi.org/10.1037/h0048826

Gibson JJ. (1979) **The ecological approach to visual perception**

Gold JI, Stocker AA (2017) **Visual Decision-Making in an Uncertain and Dynamic World** *Annual Review of Vision Science* **9**:227–250 https://doi.org/10.1146/annurev-vision-111815-114511

Gold JI, Watanabe T (2010) **Perceptual learning** *Current biology : CB* **20** https://doi.org/10.1016/j.cub.2009.10.066

Hochstein S, Ahissar M (2002) **View from the Top Hierarchies and Reverse Hierarchies in the Visual System** *Neuron* **36**:791–804 https://doi.org/10.1016/S0896-6273(02)01091-7

Hua T, Bao P, Huang CB, Wang Z, Xu J, Zhou Y, Lu ZL (2010) **Perceptual learning improves contrast sensitivity of V1 neurons in cats** *Current biology: CB* **20**:887–894 https://doi.org/10.1016/j.cub.2010.03.066

Hung SC, Seitz AR (2014) **Prolonged training at threshold promotes robust retinotopic specificity in perceptual learning** *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* **34**:8423–8431 https://doi.org/10.1523/JNEUROSCI.0745-14.2014

Jeter PE, Dosher BA, Petrov A, Lu ZL (2009) **Task precision at transfer determines specificity of perceptual learning** *Journal of Vision* **9**:1–1 https://doi.org/10.1167/9.3.1

Klein F (1893) **Vergleichende Betrachtungen über neuere geometrische Forschungen** *Mathematische Annalen* **43**:63–100 https://doi.org/10.1007/BF01446615

Klein F (1941) **Elementary Mathematics from an Advanced Standpoint: Geometry. Washington** *D.C.: National Mathematics Magazine*

Krizhevsky A, Sutskever I, Hinton GE (2017) **ImageNet classification with deep convolutional neural networks** *Communications of the ACM* **60**:84–90 https://doi.org/10.1145/3065386

Liu Z, Weinshall D (2000) **Mechanisms of generalization in perceptual learning** *Vision Research* **40**:97–109 https://doi.org/10.1016/S0042-6989(99)00140-6

Manenti GL, Dizaji AS, Schwiedrzik CM (2023) **Variability in training unlocks generalization in visual perceptual learning through invariant representations** *Current Biology* **1** https://doi.org/10.1016/j.cub.2023.01.011

Markov NT, Ercsey-Ravasz M, Essen DCV, Knoblauch K, Toroczkai Z, Kennedy H (2013) **Cortical High-Density Counterstream Architectures** *Science* **342** https://doi.org/10.1126/science.123840

Mascetti L *et al.* (2013) **The Impact of Visual Perceptual Learning on Sleep and Local Slow-Wave Initiation** *The Journal of Neuroscience* **33**:3323–3331 https://doi.org/10.1523/JNEUROSCI.0763-12.2013

McGovern DP, Webb BS, Peirce JW (2012) **Transfer of perceptual learning between different visual tasks** *Journal of Vision* **12** https://doi.org/10.1167/12.11.4

Meng Q, Wang B, Cui D, Liu N, Huang YB, Chen L, Ma Y (2019) **Age-related changes in local and global visual perception** *Journal of vision* **19** https://doi.org/10.1167/19.1.10

Orsten-Hooge K, Portillo M, Pomerantz J (2011) **False Pop Out in Visual Search** *Journal of Vision* **9**:1314–1314 https://doi.org/10.1167/11.11.1314

Pelli DG (1997) **The VideoToolbox software for visual psychophysics: transforming numbers into movies** *Spatial Vision* **10**:437–442 https://doi.org/10.1163/156856897X00366

Petrov AA, Van Horn NM, Ratcliff R (2011) **Dissociable perceptual-learning mechanisms revealed by diffusion-model analysis** *Psychonomic Bulletin & Review* **18**:490–497 https://doi.org/10.3758/s13423-011-0079-8

Sowden PT, Rose D, Davies IRL (2002) **Perceptual learning of luminance contrast detection: specific for spatial frequency and retinal location but not orientation** *Vision Research* **42**:1249–1258 https://doi.org/10.1016/s0042-6989(02)00019-6

Szokolszky A, Read CY, Palatinus Z, Palatinus K (2019) **Ecological approaches to perceptual learning: learning to perceive and perceiving as learning** *Adaptive Behavior* **27**:363–388 https://doi.org/10.1177/1059712319854687

Szpiro SFA, Carrasco M (2015) **Exogenous Attention Enables Perceptual Learning** *Psychological Science* **26**:1854–1862 https://doi.org/10.1177/0956797615598976

Todd JT, Weismantel E, Kallie CS (2014) **On the relative detectability of configural properties** *Journal of Vision* **14**:18–18 https://doi.org/10.1167/14.1.18

Todd JT, Chen L, Norman JF (1998) **On the Relative Salience of Euclidean, Affine, and Topological Structure for 3-D Form Discrimination** *Perception* **27**:273–282 https://doi.org/10.1068/p270273

Watson AB, Pelli DG (1983) **Quest: A Bayesian adaptive psychometric method** *Perception & Psychophysics* **33**:113–120 https://doi.org/10.3758/BF03202828

Wenliang LK, Seitz AR (2018) **Deep Neural Networks for Modeling Visual Perceptual Learning** *The Journal of Neuroscience* **38**:6028–6044 https://doi.org/10.1523/JNEUROSCI.1620-17.2018

Yang J *et al.* (2022) **Identifying Long- and Short-Term Processes in Perceptual Learning** *Psychological Science* **33**:830–843 https://doi.org/10.1177/09567976211056620

Zhang JY, Zhang GL, Xiao LQ, Klein SA, Levi DM, Yu C (2010) **Rule-Based Learning Explains Visual Perceptual Learning and Its Specificity and Transfer** *Journal of Neuroscience* **30**:12323–12328 https://doi.org/10.1523/JNEUROSCI.0704-10.2010

Article and author information

Yan Yang

State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing 100049, China
ORCID iD: 0009-0003-4429-8443

Yan Zhuo
State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China

Zhentao Zuo
State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China, Hefei Comprehensive National Science Center, Institute of Artificial Intelligence, Hefei 230088, China, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing 100049, China
For correspondence: zuozt@ibp.ac.cn
ORCID iD: 0000-0002-5909-3884

Tiangang Zhuo
State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China, Hefei Comprehensive National Science Center, Institute of Artificial Intelligence, Hefei 230088, China, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing 100049, China
For correspondence: zuozt@ibp.ac.cn
ORCID iD: 0009-0005-1005-5106

Lin Chen
State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China, Hefei Comprehensive National Science Center, Institute of Artificial Intelligence, Hefei 230088, China, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing 100049, China

Copyright

© 2024, Yang et al.

This article is distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use and redistribution provided that the original author and source are credited.

Editors

Reviewing Editor
SP Arun
Indian Institute of Science Bangalore, Bangalore, India

Senior Editor
Floris de Lange
Donders Institute for Brain, Cognition and Behaviour, Nijmegen, Netherlands

Reviewer #1 (Public Review):

Summary:
Visual Perceptual Learning (VPL) results in varying degrees of generalization to tasks or stimuli not seen during training. The question of which stimulus or task features predict whether learning will transfer to a different perceptual task has long been central in the field of perceptual learning, with numerous theories proposed to address it. This paper introduces a novel framework for understanding generalization in VPL, focusing on the form invariants of the training stimulus. Contrary to a previously proposed theory that task difficulty predicts

the extent of generalization - suggesting that more challenging tasks yield less transfer to other tasks or stimuli - this paper offers an alternative perspective. It introduces the concept of task invariants and investigates how the structural stability of these invariants affects VPL and its generalization. The study finds that tasks with high-stability invariants are learned more quickly. However, training with low-stability invariants leads to greater generalization to tasks with higher stability, but not the reverse. This indicates that, at least based on the experiments in this paper, an easier training task results in less generalization, challenging previous theories that focus on task difficulty (or precision). Instead, this paper posits that the structural stability of stimulus or task invariants is the key factor in explaining VPL generalization across different tasks

Strengths:
- The paper effectively demonstrates that the difficulty of a perceptual task does not necessarily correlate with its learning generalization to other tasks, challenging previous theories in the field of Visual Perceptual Learning. Instead, it proposes a significant and novel approach, suggesting that the form invariants of training stimuli are more reliable predictors of learning generalization. The results consistently bolster this theory, underlining the role of invariant stability in forecasting the extent of VPL generalization across different tasks.

- The experiments conducted in the study are thoughtfully designed and provide robust support for the central claim about the significance of form invariants in VPL generalization.

Weaknesses:
- The paper assumes a considerable familiarity with the Erlangen program and the definitions of invariants and their structural stability, potentially alienating readers who are not versed in these concepts. This assumption may hinder the understanding of the paper's theoretical rationale and the selection of stimuli for the experiments, particularly for those unfamiliar with the Erlangen program's application in psychophysics. A brief introduction to these key concepts would greatly enhance the paper's accessibility. The justification for the chosen stimuli and the design of the three experiments could be more thoroughly articulated.

- The paper does not clearly articulate how its proposed theory can be integrated with existing observations in the field of VPL. While it acknowledges previous theories on VPL generalization, the paper falls short in explaining how its framework might apply to classical tasks and stimuli that have been widely used in the VPL literature, such as orientation or motion discrimination with Gabors, vernier acuity, etc. It also does not provide insight into the application of this framework to more naturalistic tasks or stimuli. If the stability of invariants is a key factor in predicting a task's generalization potential, the paper should elucidate how to define the stability of new stimuli or tasks. This issue ties back to the earlier mentioned weakness: namely, the absence of a clear explanation of the Erlangen program and its relevant concepts.

- The paper does not convincingly establish the necessity of its introduced concept of invariant stability for interpreting the presented data. For instance, consider an alternative explanation: performing in the collinearity task requires orientation invariance. Therefore, it's straightforward that learning the collinearity task doesn't aid in performing the other two tasks (parallelism and orientation), which do require orientation estimation. Interestingly, orientation invariance is more characteristic of higher visual areas, which, consistent with the Reverse Hierarchy Theory, are engaged more rapidly in learning compared to lower visual areas. This simpler explanation, grounded in established concepts of VPL and the tuning properties of neurons across the visual cortex, can account for the observed effects, at least in one scenario. This approach has previously been used/proposed to explain VPL generalization, as seen in (Chowdhury and DeAngelis, Neuron, 2008), (Liu and Pack, Neuron, 2017), and (Bakhtiari et al., JoV, 2020). The question then is: how does the concept of invariant stability provide additional insights beyond this simpler explanation?

- While the paper discusses the transfer of learning between tasks with varying levels of invariant stability, the mechanism of this transfer within each invariant condition remains unclear. A more detailed analysis would involve keeping the invariant's stability constant while altering a feature of the stimulus in the test condition. For example, in the VPL literature, one of the primary methods for testing generalization is examining transfer to a new stimulus location. The paper does not address the expected outcomes of location transfer in relation to the stability of the invariant. Moreover, in the affine and Euclidean conditions one could maintain consistent orientations for the distractors and targets during training, then switch them in the testing phase to assess transfer within the same level of invariant structural stability.

- In the section detailing the modeling experiment using deep neural networks (DNN), the takeaway was unclear. While it was interesting to observe that the DNN exhibited a generalization pattern across conditions similar to that seen in the human experiments, the claim made in the abstract and introduction that the model provides a 'mechanistic' explanation for the phenomenon seems overstated. The pattern of weight changes across layers, as depicted in Figure 7, does not conclusively explain the observed variability in generalizations. Furthermore, the substantial weight change observed in the first two layers during the orientation discrimination task is somewhat counterintuitive. Given that neurons in early layers typically have smaller receptive fields and narrower tunings, one would expect this to result in less transfer, not more.

https://doi.org/10.7554/eLife.93959.1.sa1

Reviewer #2 (Public Review):

The strengths of this paper are clear: The authors are asking a novel question about geometric representation that would be relevant to a broad audience. Their question has a clear grounding in pre-existing mathematical concepts, that, to my knowledge, have been only minimally explored in cognitive science. Moreover, the data themselves are quite striking, such that my only concern would be that the data seem almost *too* clean. It is hard to know what to make of that, however. From one perspective, this is even more reason the results should be publicly available. Yet I am of the (perhaps unorthodox) opinion that reviewers should voice these gut reactions, even if it does not influence the evaluation otherwise. Below I offer some more concrete comments:

(1) The justification for the designs is not well explained. The authors simply tell the audience in a single sentence that they test projective, affine, and Euclidean geometry. But despite my familiarity with these terms -- familiarity that many readers may not have -- I still had to pause for a very long time to make sense of how these considerations led to the stimuli that were created. I think the authors must, for a point that is so central to the paper, thoroughly explain exactly why the stimuli were designed the way that they were and how these designs map onto the theoretical constructs being tested.

(2) I wondered if the design in Experiment 1 was flawed in one small but critical way. The goal of the parallelism stimuli, I gathered, was to have a set of items that is not parallel to the other set of items. But in doing that, isn't the manipulation effectively the same as the manipulation in the orientation stimuli? Both functionally involve just rotating one set by a fixed amount. (Note: This does not seem to be a problem in Experiment 2, in which the conditions are more clearly delineated.)

(3) I wondered if the results would hold up for stimuli that were more diverse. It seems that a determined experimenter could easily design an "adversarial" version of these experiments for which the results would be unlikely to replicate. For instance: In the orientation group in

Experiment 1, what if the odd-one-out was rotated 90 degrees instead of 180 degrees? Intuitively, it seems like this trial type would now be much easier, and the pattern observed here would not hold up. If it did hold up, that would provide stronger support for the authors' theory.

It is not enough, in my opinion, to simply have some confirmatory evidence of this theory. One would have to have thoroughly tested many possible ways that theory could fail. I'm unsure that enough has been done here to convince me that these ideas would hold up across a more diverse set of stimuli.

https://doi.org/10.7554/eLife.93959.1.sa0

Author Response

Reviewer #1 (Public Review):

Summary:

Visual Perceptual Learning (VPL) results in varying degrees of generalization to tasks or stimuli not seen during training. The question of which stimulus or task features predict whether learning will transfer to a different perceptual task has long been central in the field of perceptual learning, with numerous theories proposed to address it. This paper introduces a novel framework for understanding generalization in VPL, focusing on the form invariants of the training stimulus. Contrary to a previously proposed theory that task difficulty predicts the extent of generalization - suggesting that more challenging tasks yield less transfer to other tasks or stimuli - this paper offers an alternative perspective. It introduces the concept of task invariants and investigates how the structural stability of these invariants affects VPL and its generalization. The study finds that tasks with high-stability invariants are learned more quickly. However, training with low-stability invariants leads to greater generalization to tasks with higher stability, but not the reverse. This indicates that, at least based on the experiments in this paper, an easier training task results in less generalization, challenging previous theories that focus on task difficulty (or precision). Instead, this paper posits that the structural stability of stimulus or task invariants is the key factor in explaining VPL generalization across different tasks

Strengths:

- *The paper effectively demonstrates that the difficulty of a perceptual task does not necessarily correlate with its learning generalization to other tasks, challenging previous theories in the field of Visual Perceptual Learning. Instead, it proposes a significant and novel approach, suggesting that the form invariants of training stimuli are more reliable predictors of learning generalization. The results consistently bolster this theory, underlining the role of invariant stability in forecasting the extent of VPL generalization across different tasks.*

- *The experiments conducted in the study are thoughtfully designed and provide robust support for the central claim about the significance of form invariants in VPL generalization.*

Weaknesses:

- *The paper assumes a considerable familiarity with the Erlangen program and the definitions of invariants and their structural stability, potentially alienating*

> *readers who are not versed in these concepts. This assumption may hinder the understanding of the paper's theoretical rationale and the selection of stimuli for the experiments, particularly for those unfamiliar with the Erlangen program's application in psychophysics. A brief introduction to these key concepts would greatly enhance the paper's accessibility. The justification for the chosen stimuli and the design of the three experiments could be more thoroughly articulated.*

Response: We appreciate the reviewer's feedback regarding the accessibility of our paper. In response to this feedback, we plan to enhance the introduction section of our paper to provide a concise yet comprehensive overview of the key concepts of Erlangen program. Additionally, we will provide a more thorough justification for the selection of stimuli and the experimental design in our revised version, ensuring that readers understand the rationale behind our choices.

> - *The paper does not clearly articulate how its proposed theory can be integrated with existing observations in the field of VPL. While it acknowledges previous theories on VPL generalization, the paper falls short in explaining how its framework might apply to classical tasks and stimuli that have been widely used in the VPL literature, such as orientation or motion discrimination with Gabors, vernier acuity, etc. It also does not provide insight into the application of this framework to more naturalistic tasks or stimuli. If the stability of invariants is a key factor in predicting a task's generalization potential, the paper should elucidate how to define the stability of new stimuli or tasks. This issue ties back to the earlier mentioned weakness: namely, the absence of a clear explanation of the Erlangen program and its relevant concepts.*

Response: Thanks for highlighting the need for better integration of our proposed theory with existing observations in the field of VPL. Unfortunately, the theoretical framework proposed in our study is based on the Klein's Erlangen program and is only applicable to geometric shape stimuli. For VPL studies using stimuli and paradigms that are completely unrelated to geometric transformations (such as motion discrimination with Gabors or random dots, vernier acuity, spatial frequency discrimination, contrast detection or discrimination, etc.), our proposed theory does not apply. Some stimuli employed by VPL studies can be classified into certain geometric invariants. For instance, orientation discrimination with Gabors (Dosher & Lu, 2005) and texture discrimination task (F. Wang et al., 2016) both belong to tasks involving Euclidean invariants, and circle versus square discrimination (Kraft et al., 2010) belongs to tasks involving affine invariance. However, these studies do not simultaneously involve multiple geometric invariants of varying levels stability, and thus cannot be directly compared with our research. It is worth noting that while the Klein's hierarchy of geometries, which our study focuses on, is rarely mentioned in the field of VPL, it does have connections with concepts such as 'global/local', 'coarse/fine', 'easy/difficulty', 'complex/simple': more stable invariants are closer to 'global', 'coarse', 'easy', 'complex', while less stable invariants are closer to 'local', 'fine', 'difficulty', 'simple'. Importantly, several VPL studies have found 'fine-to-coarse' or 'local-to-global' asymmetric transfer (Chang et al., 2014; N. Chen et al., 2016; Dosher & Lu, 2005), which seems consistent with the results of our study.

In the introduction section of our revised version and subsequent full author response, we will provide a clear explanation of the Erlangen program and elucidate how to define the stability of new stimuli or tasks. In the discussion section of our revised version, we will compare our results to other studies concerned with the generalization of perceptual learning and speculate on how our proposed theory fit with existing observations in the field of VPL.

- *The paper does not convincingly establish the necessity of its introduced concept of invariant stability for interpreting the presented data. For instance, consider an alternative explanation: performing in the collinearity task requires orientation invariance. Therefore, it's straightforward that learning the collinearity task doesn't aid in performing the other two tasks (parallelism and orientation), which do require orientation estimation. Interestingly, orientation invariance is more characteristic of higher visual areas, which, consistent with the Reverse Hierarchy Theory, are engaged more rapidly in learning compared to lower visual areas. This simpler explanation, grounded in established concepts of VPL and the tuning properties of neurons across the visual cortex, can account for the observed effects, at least in one scenario. This approach has previously been used/proposed to explain VPL generalization, as seen in (Chowdhury and DeAngelis, Neuron, 2008), (Liu and Pack, Neuron, 2017), and (Bakhtiari et al., JoV, 2020). The question then is: how does the concept of invariant stability provide additional insights beyond this simpler explanation?*

Response: We appreciate the alternative explanation proposed by the reviewer and agree that it presents a valid perspective grounded in established concepts of VPL and neural tuning properties. However, performing in the collinearity and parallelism tasks both require orientation invariance. While utilizing the orientation invariance, as proposed by the reviewer, can explain the lack of transfer from collinearity or parallelism to orientation task, it cannot explain why collinearity does not transfer to parallelism.

As stated in the response to the previous review, in the revised discussion section, we will compare our study with other studies (including the three papers mentioned by the reviewer), aiming to clarify the necessity of the concept of invariant stability for interpreting the observed data and understanding the mechanisms underlying VPL generalization.

- *While the paper discusses the transfer of learning between tasks with varying levels of invariant stability, the mechanism of this transfer within each invariant condition remains unclear. A more detailed analysis would involve keeping the invariant's stability constant while altering a feature of the stimulus in the test condition. For example, in the VPL literature, one of the primary methods for testing generalization is examining transfer to a new stimulus location. The paper does not address the expected outcomes of location transfer in relation to the stability of the invariant. Moreover, in the affine and Euclidean conditions one could maintain consistent orientations for the distractors and targets during training, then switch them in the testing phase to assess transfer within the same level of invariant structural stability.*

Response: Thanks for raising the issue regarding the mechanism of transfer within each invariant conditions. We plan to design an additional experiment that is similar in paradigm to Experiment 2, aiming to examine how VPL generalizes to a new test location within a single invariant stability level.

- *In the section detailing the modeling experiment using deep neural networks (DNN), the takeaway was unclear. While it was interesting to observe that the DNN exhibited a generalization pattern across conditions similar to that seen in the human experiments, the claim made in the abstract and introduction that the model provides a 'mechanistic' explanation for the phenomenon seems overstated. The pattern of weight changes across layers, as depicted in Figure 7, does not conclusively explain the observed variability in generalizations. Furthermore, the substantial weight change observed in the first two layers*

> *during the orientation discrimination task is somewhat counterintuitive. Given that neurons in early layers typically have smaller receptive fields and narrower tunings, one would expect this to result in less transfer, not more.*

Response: We appreciate the reviewer's feedback regarding the clarity of our DNN modeling experiment. We acknowledge that while DNNs have been demonstrated to serve as models for visual systems as well as VPL, the claim that the model provides a 'mechanistic' explanation for the phenomenon still overstated. In our revised version,

We will attempt a more detailed analysis of the DNN model while providing a more explicit explanation of the findings from the DNN modeling experiment, emphasizing its implications for understanding the observed variability in generalizations.

Additionally, the substantial weight change observed in the first two layers during the orientation discrimination task is not contradictory to the theoretical framework we proposed, instead, it aligns with our speculation regarding the neural mechanisms of VPL for geometric invariants. Specifically, it suggests that invariants with lower stability rely more on the plasticity of lower-level brain areas, thus exhibiting poorer generalization performance to new locations or stimulus features within each invariant conditions. However, it does not imply that their learning effects cannot transfer to invariants with higher stability.

> ***Reviewer #2 (Public Review):***
>
> *The strengths of this paper are clear: The authors are asking a novel question about geometric representation that would be relevant to a broad audience. Their question has a clear grounding in pre-existing mathematical concepts, that, to my knowledge, have been only minimally explored in cognitive science. Moreover, the data themselves are quite striking, such that my only concern would be that the data seem almost* too *clean. It is hard to know what to make of that, however. From one perspective, this is even more reason the results should be publicly available. Yet I am of the (perhaps unorthodox) opinion that reviewers should voice these gut reactions, even if it does not influence the evaluation otherwise. Below I offer some more concrete comments:*
>
> *(1) The justification for the designs is not well explained. The authors simply tell the audience in a single sentence that they test projective, affine, and Euclidean geometry. But despite my familiarity with these terms -- familiarity that many readers may not have -- I still had to pause for a very long time to make sense of how these considerations led to the stimuli that were created. I think the authors must, for a point that is so central to the paper, thoroughly explain exactly why the stimuli were designed the way that they were and how these designs map onto the theoretical constructs being tested.*
>
> *(2) I wondered if the design in Experiment 1 was flawed in one small but critical way. The goal of the parallelism stimuli, I gathered, was to have a set of items that is not parallel to the other set of items. But in doing that, isn't the manipulation effectively the same as the manipulation in the orientation stimuli? Both functionally involve just rotating one set by a fixed amount. (Note: This does not seem to be a problem in Experiment 2, in which the conditions are more clearly delineated.)*
>
> *(3) I wondered if the results would hold up for stimuli that were more diverse. It seems that a determined experimenter could easily design an "adversarial" version of these experiments for which the results would be unlikely to replicate. For instance: In the orientation group in Experiment 1, what if the odd-one-out was rotated 90 degrees instead of 180 degrees? Intuitively, it seems like this trial type would now be much easier, and the pattern observed here would not hold up. If it did hold up, that would provide stronger support for the authors' theory.*

> *It is not enough, in my opinion, to simply have some confirmatory evidence of this theory. One would have to have thoroughly tested many possible ways that theory could fail. I'm unsure that enough has been done here to convince me that these ideas would hold up across a more diverse set of stimuli.*

Response: (1) We appreciate the reviewer's feedback regarding the justification for our experimental designs. We recognize the importance of thoroughly explaining how our stimuli were designed and how these designs correspond to the theoretical constructs being tested. In our revised version, we will enhance the introduction of Erlangen program and provide a more detailed explanation of the rationale behind our stimulus designs, aiming to enhance the clarity and transparency of our experimental approach for readers who may not be familiar with these concepts.

(2) We appreciate the reviewer's insight into the design of Experiment 1 and the concern regarding the potential similarity between the parallelism and orientation stimuli manipulations.

The parallelism and orientation stimuli in Experiment 1 were first used by Olson & Attneave (1970) to support line-based models of shape coding and then adapted to measure the relative salience of different geometric properties (Chen, 1986). In the parallelism stimuli, the odd quadrant differs from the rest in line slope, while in the orientation stimuli, in contrast, the odd quadrant contains exactly the same line segments as the rest but differs in direction pointed by the angles. The result, that the odd quadrant was detected much faster in the parallelism stimuli than in the orientation stimuli, can serve as evidence for line-based models of shape coding. However, according to Chen (1986, 2005), the idea of invariants over transformations suggests a new analysis of the data: in the parallelism stimuli, the fact that line segments share the same slope essentially implies that they are parallel, and the discrimination may be actually based on parallelism. Thus, the faster discrimination of the parallelism stimuli than that of the orientation stimuli may be explained in terms of relative superiority of parallelism over orientation of angles—a Euclidean property.

The group of stimuli in Experiment 1 has been employed by several studies to investigate scientific questions related to the Klein's hierarchy of geometries (L. Chen, 2005; Meng et al., 2019; B. Wang et al., n.d.). Due to historical inheritance, we adopted this set of stimuli and corresponding paradigm, despite their imperfect design.

(3) Thanks for raising the important issue of stimulus diversity and the potential for "adversarial" versions of the experiments to challenge our findings. We acknowledge the validity of your concern and recognize the need to demonstrate the robustness of our results across a range of stimuli. We plan to design additional experiments to investigate the potential implications of varying stimulus characteristics, such as different rotation angles proposed by the reviewer, on the observed patterns of performance.