Transfer in motion discrimination learning was no greater in double training than in single training

Jinfeng Huang*
Chinese Academy of Sciences Key Laboratory of Brain Function and Disease and School of Life Sciences, University of Science and Technology of China, Hefei, China

Ju Liang*
Chinese Academy of Sciences Key Laboratory of Brain Function and Disease and School of Life Sciences, University of Science and Technology of China, Hefei, China

Yifeng Zhou
Chinese Academy of Sciences Key Laboratory of Brain Function and Disease and School of Life Sciences, University of Science and Technology of China, Hefei, China

Zili Liu
Department of Psychology, University of California Los Angeles, USA

We investigated the controversy regarding double training in motion discrimination learning. We collected data from 43 participants in a motion direction discrimination learning task with either double training (i.e., training plus exposure) or single training (i.e., no exposure). By pooling these data with those in the literature, we had data in double training from 28 participants and in single training from 36 participants. We found that, in double training, the transfer along the exposed direction was less than that along the trained direction, indicating incomplete transfer. Importantly, the transfer in double training was not reliably greater than that in single training.

Introduction

Double training (or training plus exposure) is a recent technique in perceptual learning that has been claimed to enable a degree of transfer greater than the amount of transfer prior methods can obtain (Xiao et al., 2008). The key manipulation of this technique is to pair a training task at one stimulus attribute (e.g., at one retinal location) with an irrelevant task at another attribute (e.g., at a different retinal location). As a result of such double training, perceptual learning was found to transfer from the first retinal location to the second location much more than when only single training of the first task was used. This double training technique has been applied successfully to a variety of perceptual learning tasks including, to name just a few, contrast discrimination (Xiao et al., 2008), Vernier acuity (Wang, Cong, & Yu, 2013; Wang, Zhang, Klein, Levi, & Yu, 2012, 2014), orientation discrimination (Zhang et al., 2010), visuomotor learning (Yin, Bi, Yu, & Wei, 2016), and motion direction discrimination (Zhang & Yang, 2014).

The main message from these double training studies has been that the signature finding of stimulus specificity in traditional perceptual learning may not be fundamental after all. Instead, stimulus exposure at a new attribute (e.g., at a new retinal location) sufficed to enable transfer to this new location, where little transfer had been found without the exposure. In fact, Zhang et al. (2010) argued that perceptual learning was in principle little different from higher-level rule-based learning. Given that visual perceptual learning had been hypothesized to be largely, although not entirely, low level that takes place in early visual cortices (for reviews, see Fahle, 2005; Gilbert, 1994; Sagi, 2011; Watanabe & Sasaki, 2015), the double training proposal offers a completely different perspective on

Citation: Huang, J., Liang, J., Zhou, Y., & Liu, Z. (2017). Transfer in motion discrimination learning was no greater in double training than in single training. *Journal of Vision*, *17*(6):7, 1–10, doi:10.1167/17.6.7.

doi: 10.1167/17.6.7 Received May 4, 2016; published June 14, 2017 ISSN 1534-7362

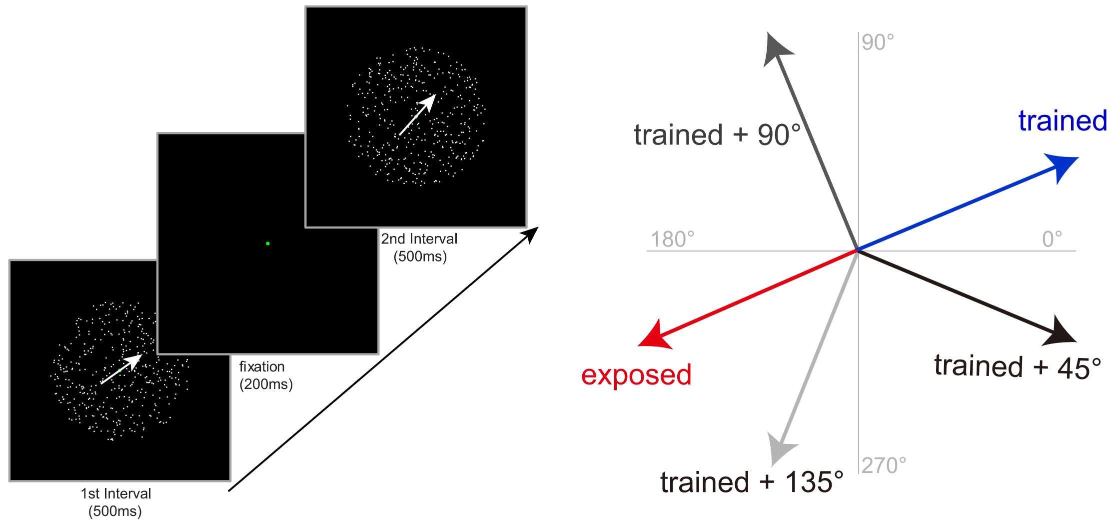

Figure 1. (Left) Schematic illustration of a single trial in motion discrimination. In the first task, participants decided whether the relative change of direction was clockwise or counterclockwise. In the second task, the dots moved along directions opposite to those in the first task. Participants decided whether the first or second stimulus had more dots. (Right) The five reference directions used in the experiments. The "trained" direction refers to the average direction in motion discrimination training; the "exposed" refers to the second training direction on numerosity. These two directions were counterbalanced. The remaining three directions were control directions. In double training, the two tasks were trained in alternating blocks for five daily sessions. At pre- and posttraining, motion discrimination thresholds were measured along all five reference directions. In single training, everything was kept the same, except that the second task was not run.

what might be going on in perceptual learning. If proven correct, its impact to the field of perceptual learning will be highly significant. Because of this significance, we decided to verify this technique in motion direction discrimination learning, with which we had had some experience.

Before specifying the hypotheses that were tested in the current study, it is useful to describe in detail double training in motion discrimination. In Zhang and Yang (2014), six participants were shown two sequential random-dot motion stimuli and decided whether the motion direction changed clockwise or counterclockwise (Figure 1). This was the first training task. In the second task, two random-dot motion stimuli were sequentially shown in a new average direction opposite to that in the first task. The participants decided whether the dot number was increased or decreased. These two tasks were trained in alternating blocks. After the double training, motion discrimination was remeasured along all test directions and was found to transfer completely from the first to the second task direction. The amount of transfer, defined as transfer index (TI; TI = 1 – posttraining threshold/pretraining threshold) was 120%. The six participants' indices were 78%, 182%, 44%, 270%, 62%, and 87%.

To establish that this complete transfer of 120% was indeed due to the second task, Zhang and Yang (2014) created the following two control conditions. The first control was within-subject, in which transfer was measured along directions that were neither trained nor exposed. The TI was found to be 47%. The second control was between-subjects, in which the same experiment was repeated with six new participants, except that there was no second task. The amount of transfer from this single training experiment was termed *baseline transfer* and was found to be TI = 17% (the six indices were −10%, 0%, 6%, 9%, 97%, and 0%). Based on these control comparisons, Zhang and Yang (2014) concluded that double training enabled substantially greater transfer than single training.

In an effort to verify this double training technique, Liang, Zhou, Fahle, and Liu (2015a) first set out to extend the technique to long-term learning (48 sessions) using the method of constant stimuli, in which only two fixed directions were used in motion discrimination. They found very limited transfer, however. As a result, Liang, Zhou, Fahle, and Liu (2015b) decided to replicate the Zhang and Yang (2014) study with the method of staircase. There, the resultant double training transfer (TI = 48%, $N = 9$) was substantially smaller than that in Zhang and Yang (2014; TI = 120%, $N = 6$). Liang et al. (2015b) also found little difference between the transfers along the exposed and nonexposed directions.

In response to Liang et al. (2015a), Xiong, Xie, and Yu (2016) conducted double training in motion discrimination with constant stimuli (seven sessions) and confirmed that there was little transfer. They then

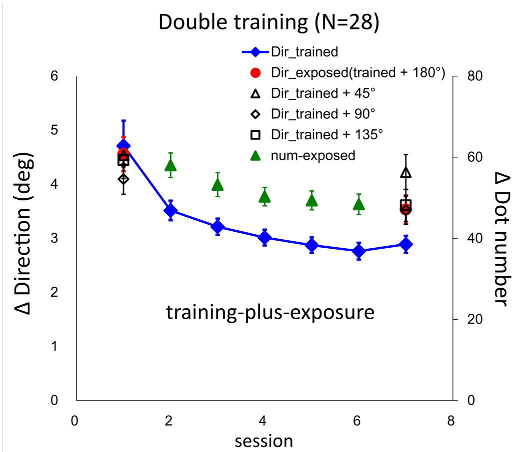

Figure 2. Double training: motion discrimination thresholds (in blue, red, and black) and numerosity discrimination thresholds (in green) from 28 available participants. The participants were double trained with motion discrimination along one reference direction (blue) and with numerosity discrimination along the opposite reference direction (green). At pre- (Session 1) and posttraining (Session 7), motion discrimination thresholds were measured along the trained direction (blue), second task (or exposed) direction (red), and three control directions (black). Double training would predict a greater drop of threshold along the exposed direction than when there was only single training (or no exposure). The *y*-axis on the right shows the numerosity discrimination performance. Error bars represent standard errors of the mean.

jittered the directions within a range of ±2° while keeping the angular difference between the two directions constant. With the jittering, they found substantially more transfer. In the same study, Xiong et al. (2016) also conducted double training in motion discrimination with staircase, similar to that in Zhang and Yang (2014). Critically, the baseline transfer of 17% in single training that was used to compare with double training was again from the same six participants in Zhang and Yang (2014).

Using still the same six participants' data of 17% as a baseline, Zhang and Yu (2016) pooled together 24 participants' double training data in the literature. In this analysis, the transfer to the second task direction was found to be 78%. They concluded that double training indeed enabled substantially greater transfer than single training (17%).

Given the apparent importance of an accurate measurement of the baseline transfer, and given large individual differences in perceptual learning in general, it is desirable to establish such baseline from more participants. The goal of the current study was to replicate the single and double training experiments with more participants, so that their data could be pooled with existing data to more firmly establish the effect size of the double training. We ran a double and a single training experiments with 43 participants and pooled these together with available data in the literature to test whether or not double training transferred more than single training.

Experiments

All experiments were conducted at the University of Science and Technology of China, Hefei, China. The same computer code, as well as experimental instructions and protocol, were used as in Zhang and Yang (2014). In particular, all participants practiced one block of motion discrimination (about 50 trials) prior to their experiments, as in Zhang and Yang (2014). The experimental setting was identical to that reported in Liang et al. (2015b).

Participants

Forty-three participants were recruited in accordance with the Helsinki Declaration. Thirteen were randomly assigned to participate in the double training experiment and 30 in the single training experiment. We pooled data from Zhang and Yang (2014; $N = 6$) and Liang et al. (2015b; $N = 9$) in their double training experiments, making a total of 28 participants in double training. We also pooled data from Zhang and Yang (2014; $N = 6$) in their single training experiment, making a total of 36 participants in single training. The only missing data were from Xiong et al. (2016; $N = 6$), which were unavailable. Otherwise, the double and single training groups would have a nearly matched number of participants.

Results

Figure 2 shows the threshold performance in double training from all 28 participants. The 13 participants' individual data in the current study are shown in Appendix A. We first asked whether the threshold reduction along the motion trained direction was more than that along the second task or exposed direction. We conducted a 2 directions (training vs. exposed) × 2 times (pre- vs. posttraining) analysis of variance (ANOVA). The main effect of time was significant, $F(1, 27) = 37.88$, $p = 1.41\text{E-}6$. The main effect of direction was significant, $F(1, 27) = 4.40$, $p = 0.045$. Importantly, the interaction effect was also significant, $F(1, 27) = 4.70$, $p = 0.039$, which indicated that the threshold reduction along the exposure direction (from 4.56° ± 0.32° to 3.53° ± 0.26° [$M \pm SE$]) was less than that

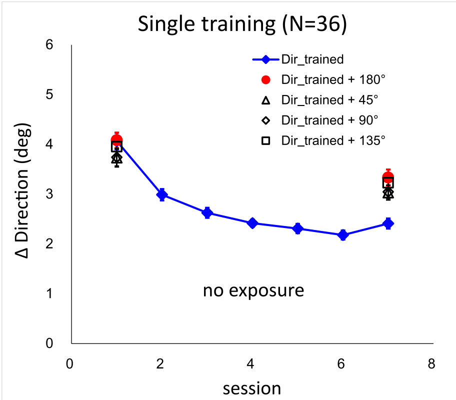

Figure 3. Single training: motion discrimination thresholds from all 36 participants in the single training experiment without the exposure task. This single training experiment was the same as the double training except there was no second task (or exposure) on numerosity discrimination. The purpose of this experiment was to establish with more participants the baseline transfer from the motion trained direction to other directions.

along the trained direction (from 4.71° ± 0.47° to 2.89° ± 0.16°). In other words, the transfer was incomplete. (The same interaction effect was obtained if we used 22 participants' data only from our own laboratories, $F[1, 21] = 5.39$, $p = 0.030$.)

Figure 3 shows the discrimination thresholds in single training from all 36 participants. To find out whether or not there was any exposure-specific effect, we conducted an ANOVA on the thresholds from both experiments. The factors were 2 experiments (double vs. single training) × 2 directions (trained vs. exposed) × 2 times (pre- vs. posttraining). The main effect of time was significant, $F(1, 62) = 134$, $p = 3.84E\text{-}17$, indicating learning as a result of training. The main effect of direction was significant, $F(1, 62) = 22$, $p = 1.42E\text{-}5$, indicating that the thresholds along the two task directions were different. The interaction between time and direction was also significant, $F(1, 62) = 22$, $p = 1.74E\text{-}5$, indicating that threshold was reduced more along the trained than that along the exposed direction. However, no other effects were significant. In particular, the interaction of Experiment × Direction × Time had an F value of only 0.10. A closer look at thresholds along the second task direction with a 2 × 2 ANOVA revealed only a significant main effect of time. The effects of experiment, $F(1, 62) = 1.29$, $p = 0.26$, and interaction, $F(1, 62) = 2.25$, $p = 0.14$, were not significant. These results suggest that there was no reliable advantage of double training over single training. Numerically, along the second task direction, double training reduced the threshold from 4.56° ± 0.32° to 3.53° ± 0.26°. In comparison, the single training reduced the thresholds from 4.08° ± 0.15° to 3.33° ± 0.16°.

The above ANOVAs were conducted on the threshold data, which had been verified not to violate the sphericity assumption. Nevertheless, in Zhang and Yang (2014) and Zhang and Yu (2016), the TI and t tests were used to quantify transfer. We therefore ran a robust Bayesian t test (Kruschke, 2013) to compare the TIs between the double and single training. No reliable difference could be found because the 95% confidence interval included zero (see Appendix B for details).

We further ran a Bayesian power analysis to estimate the number of participants needed for the current TI difference (25%) to be statistically reliable between double and single training. We found that at least 600 participants would be needed, but no more than 640, in order for the TI difference to have a power of 0.8 (Appendix B).

Discussion

The double training technique originally developed in Xiao et al. (2008) has generated significant impact in the field of perceptual learning, because substantially more transfer was found using this method than using the traditional single training method. Prior to the introduction of this technique in 2008, generalizable perceptual learning had been reported already (Ahissar & Hochstein, 1997; Harris, Gliksberg, & Sagi, 2012; Liu, 1995, 1999; Liu & Weinshall, 2000; Rubin, Nakayama, & Shapley, 1997). Nevertheless, this more recent technique of double training appeared particularly influential because the same technique has been successfully applied to many perceptual learning tasks.

In our current study, data from 43 participants were collected and pooled with existing data in the literature. The only data set not included was from the six participants in Xiong et al. (2016) because of its unavailability. The data in Liang et al. (2015a) were not usable because the method of constant stimuli was used instead of staircase, which gave rise to motion discrimination performance that was not readily translatable to discrimination thresholds. Table 1 summarizes all studies in the literature so far on motion discrimination with the double training technique.

Putting all available data together, we found that the amount of transfer to the second task direction was not reliably greater in double training than in single training. This transfer was measured in the reduction of angular threshold, which appeared well behaved as confirmed by the ANOVA sphericity test. For this reason, we believe that discrimination threshold was a better measure than TI, because TI was defined as the ratio of the improvement in the transfer direction over that in the trained direction. Yet the improvement itself was defined as the ratio of the threshold reduction over

	Publication	Number of participants	Angle between trained and exposed	Total sessions	Method
Sequential double training (the first stage was single training)	Zhang and Yang (2014)	6	180°	12	Staircase
	Current study	30	180°	7 (first stage)	Staircase
	Liang et al. (2015a)	9	135°	48	Constant stimuli
Simultaneous double training	Zhang and Yang (2014)	6	180°	7	Staircase
	Liang et al. (2015b)	9	180°	18	Staircase
	Current study	13	180°	7	Staircase
	Xiong et al. (2016)	6	90°	7	Staircase
No training	Current study	6	NA	2	Staircase

Table 1. All studies on motion discrimination learning with double training in the current study and in the literature. *Notes:* There are two versions of double training: sequential and simultaneous. In the sequential version, the first task was equivalent to single training.

the pretraining threshold. The use of this ratio of ratios produced a wide range of values (see Figure 4), including a TI as large as 251%. In an effort to compare results between using thresholds and using TIs, we also converted our data into TIs and used a robust Bayesian t test to deal with any possible "outliers." Our result using the TIs was consistent with that using the thresholds. Namely, no reliable difference between double training and single training could be found (Appendix B). Figure 4 shows individual participants' TIs from the current study ($N = 43$) and Zhang and Yang (2014; $N = 12$).

Given that no reliable difference was found between double and single training, which was a null effect, caution is needed. Numerically, the transfer as measured in threshold reduction was greater for double training (1.03°) than that for single training (0.75°). It is possible, therefore, that with an even larger number of participants, the difference could become statistically significant. However, in our Bayesian power analysis, the total number of participants required would have to be 600 at least for the TI difference found in the current study to have a power of 0.8. Hence, it appears reasonable to say that, in motion discrimination learning, there is little reason to think that there is an advantage gained in double training.

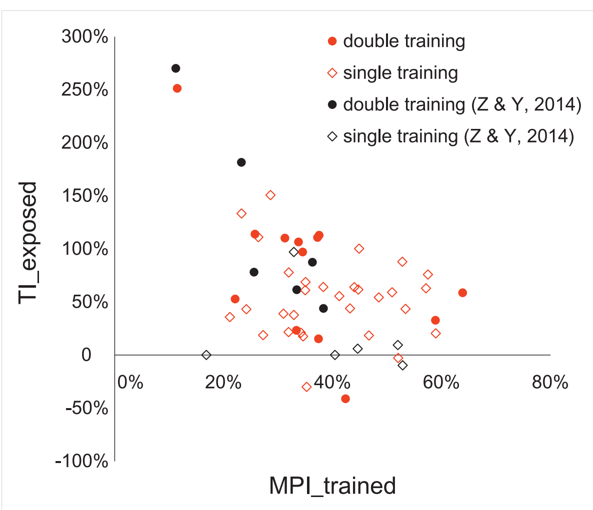

Figure 4. Individual participants' transfer indices (TIs) in their double or single training, as a function of the mean percentage improvement (MPI) along the trained direction. MPI = (1 − posttraining threshold/pretraining threshold) × 100%. The data are from the current study ($N = 13$ and 30 in double and single training, respectively) and Zhang and Yang (2014; $N = 6$ in both conditions). The two double-training participants at the top left showed transfers greater than 250%, possibly due to only modest learning.

Keywords: perceptual learning, motion perception, transfer, specificity, double training

Acknowledgments

We thank Dr. Joseph Burling for his help in the Bayesian analysis of the index data. We also thank Dr. David Bennett for his comments on the article and his editing. This research was supported in part by a fellowship to ZL at the Hanse Institute of Advanced Studies, Delmenhorst, Germany. The research was also supported in part by grants from the Natural NSF of China to ZL (NSFC 31228009) and to YZ (NSFC 31230032). We thank Dr. Jun-Yun Zhang for providing the MatLab code, experimental instructions, and the data in Zhang and Yang (2014). We also thank Drs. Jun-Yun Zhang, Stan Klein, and Cong Yu for helpful discussions.

* The first two authors contributed equally and should be considered as co-first authors.
Commercial relationships: none.
Corresponding author: Yifeng Zhou.
Email: zhouy@ustc.edu.cn.
Address: School of Life Sciences, University of Science & Technology of China, Hefei, China.

References

Ahissar, M., & Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning. *Nature*, *387*, 401–406.

Fahle, M. (2005). Perceptual learning: Specificity versus generalization. *Current Opinion in Neurobiology*, *15*, 154–160.

Gilbert, C. D. (1994). Early perceptual learning. *Proceedings of the National Academy of Science, USA*, *91*, 1195–1197.

Harris, H., Gliksberg, M., & Sagi, D. (2012). Generalized perceptual learning in the absence of sensory adaptation. *Current Biology*, *22*, 1813–1817.

Hung, S.-C., & Seitz, A. R. (2014). Prolonged training at threshold promotes robust retinotopic specificity in perceptual learning. *Journal of Neuroscience*, *34*, 8423–8431.

Kruschke, J. K. (2013). Bayesian estimation supersedes the *t* test. *Journal of Experimental Psychology: General*, *142*, 573–603. https://doi.org/10.1037/a0029146

Liang, J., Zhou, Y., Fahle, M., & Liu, Z. (2015a). Limited transfer of long-term motion perceptual learning with double training. *Journal of Vision*, *15*(10):1, 1–9, doi:10.1167/15.10.1. [PubMed] [Article]

Liang, J., Zhou, Y., Fahle, M., & Liu, Z. (2015b). Specificity of motion discrimination learning even with double training and staircase, *Journal of Vision*, *15*(10):3, 1–10, doi:10.1167/15.10.3. [PubMed] [Article]

Liu, Z. (1995). *Learning a visual skill that generalizes*. Princeton, NJ: NEC Research Institute.

Liu, Z. (1999). Perceptual learning in motion discrimination that generalizes across motion directions. *Proceedings of the National Academy of Science, USA*, *96*, 14085–14087.

Liu, Z., & Weinshall, D. (2000). Mechanisms of generalization in perceptual learning. *Vision Research*, *40*, 97–109.

Rubin, N., Nakayama, K., & Shapley, R. (1997). Abrupt learning and retinal size specificity in illusory-contour perception. *Current Biology*, *7*, 461–467.

Sagi D. (2011). Perceptual learning in vision research. *Vision Research*, *51*, 1552–1566.

Wang, R., Cong, L. J., & Yu, C. (2013). The classical TDT perceptual learning is mostly temporal learning. *Journal of Vision*, *13*(5):9, 1–9, doi:10.1167/13.5.9. [PubMed] [Article]

Wang, R., Zhang, J. Y., Klein, S. A., Levi, D. M., & Yu, C. (2012). Task relevancy and demand modulate double-training enabled transfer of perceptual learning. *Vision Research*, *61*, 33–38.

Wang, R., Zhang, J. Y., Klein, S. A., Levi, D. M., & Yu, C. (2014). Vernier perceptual learning transfers to completely untrained retinal locations after double training: A "piggybacking" effect. *Journal of Vision*, *14*(13):12, 1–12, doi:10.1167/14.13.12. [PubMed] [Article]

Watanabe, T., & Sasaki, Y. (2015) Perceptual learning: Toward a comprehensive theory. *Annual Review of Psychology*, *3*, 197–221.

Xiao, L. Q., Zhang, J. Y., Wang, R., Klein, S. A., Levi, D. M., & Yu, C. (2008). Complete transfer of perceptual learning across retinal locations enabled by double training. *Current Biology*, *18*, 1922–1926.

Xiong, Y. Z., Xie, X. Y., & Yu, C. (2016). Location and direction specificity in motion direction learning associated with a single-level method of constant stimuli. *Vision Research*, *119*, 9–15.

Yin, C., Bi, Y. Q., Yu, C., & Wei, K. L. (2016). Eliminating direction specificity in visuomotor learning. *Journal of Neuroscience*, *36*, 3839–3847.

Zhang, J. Y., & Yang, Y. X. (2014). Perceptual learning of motion direction discrimination transfers to an opposite direction with TPE training. *Vision Research*, *99*, 93–98.

Zhang, J. Y., & Yu, C. (2016). The transfer of motion direction learning to an opposite direction enabled by double training: A reply to Liang et al. (2015). *Journal of Vision*, *16*(3):29, 1–4, doi: 10.1167/16.3.29. [PubMed] [Article]

Zhang, J. Y., Zhang, G. L., Xiao, L. Q., Klein, S. A., Levi, D. M., & Yu, C. (2010). Rule-based learning explains visual perceptual learning and its specificity and transfer. *Journal of Neuroscience*, *30*, 12323–12328.

Appendix A: Data from individual participants

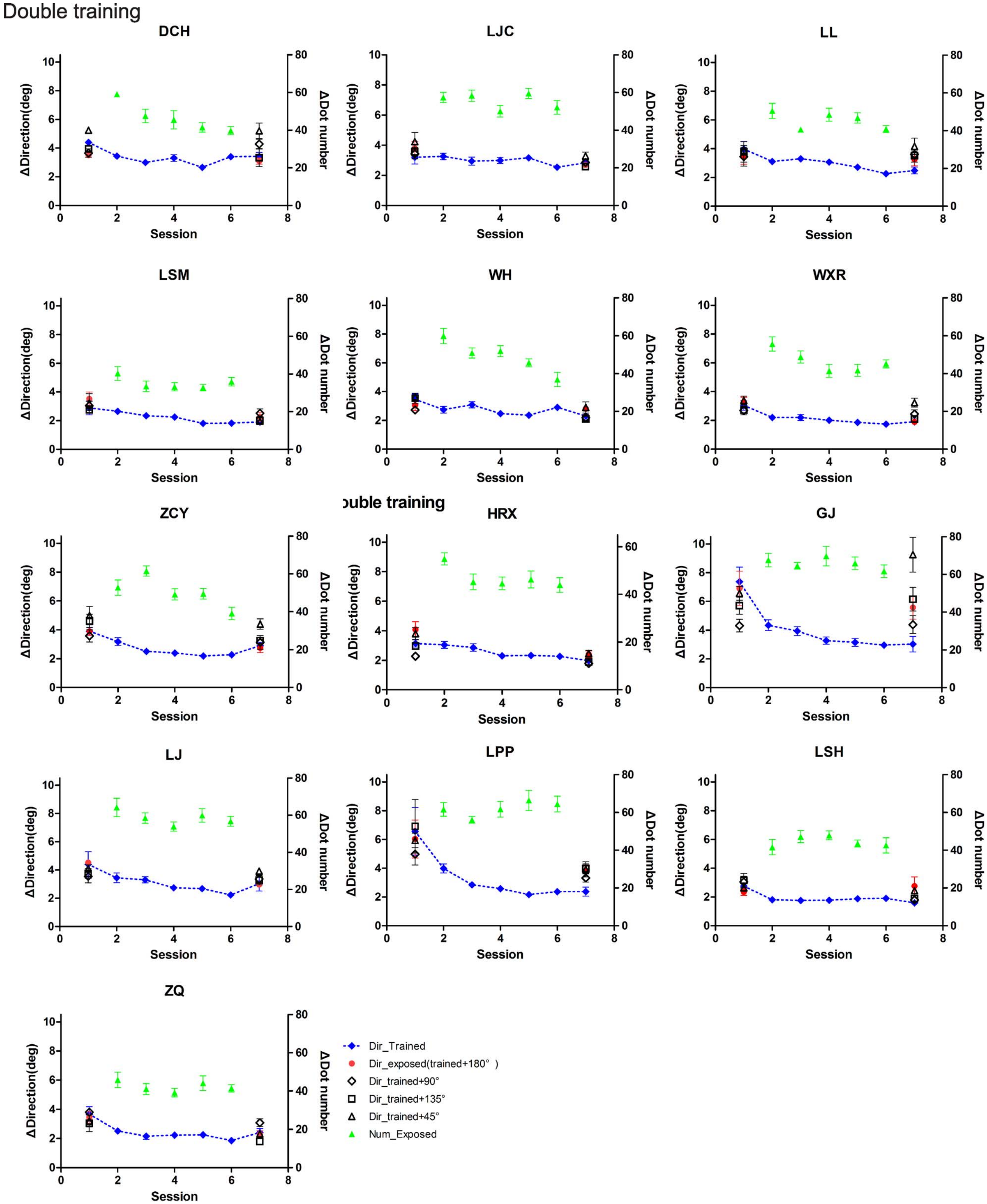

Figure 5. Double training data from each of the 13 participants in the current study. Among the remaining 15 of the 28 participants used in the current study, six were from experiment 2 in Zhang and Yang (2014) and nine were from Liang et al. (2015b). For ease of comparison, the data in this figure and those in Liang et al. (2015b) were plotted using the same color scheme and style. The only difference between these plots and those in Zhang and Yang (2014) is that the *y*-axes here and in Liang et al. (2015b) are in linear scale, whereas those in Zhang and Yang (2014) are in log scale.

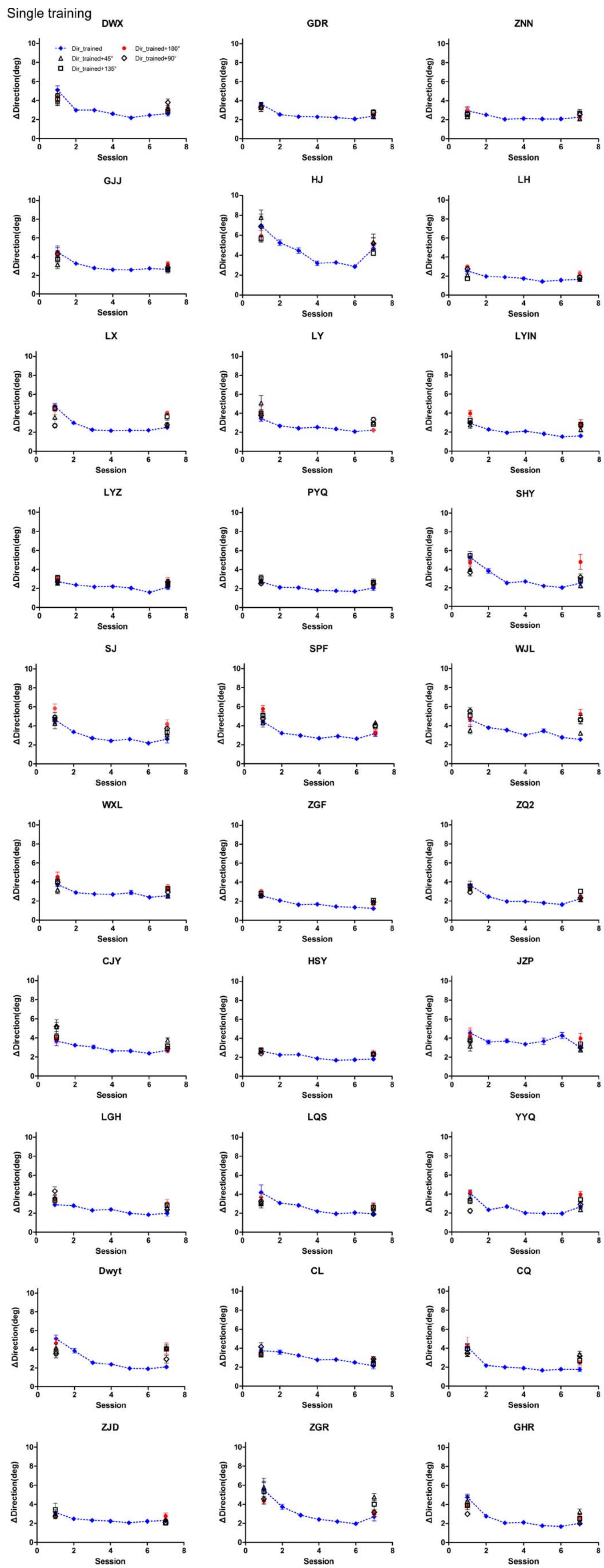

Figure 6. Single training data from each of the 30 participants in the current study. Participants SHY, ZNN, LYIN, and GDR had 1 hr of psychophysical experience that was unrelated to motion discrimination. In Figure 3, data from 36 participants were plotted. The six other participants' data were from experiment 1 in Zhang and Yang (2014).

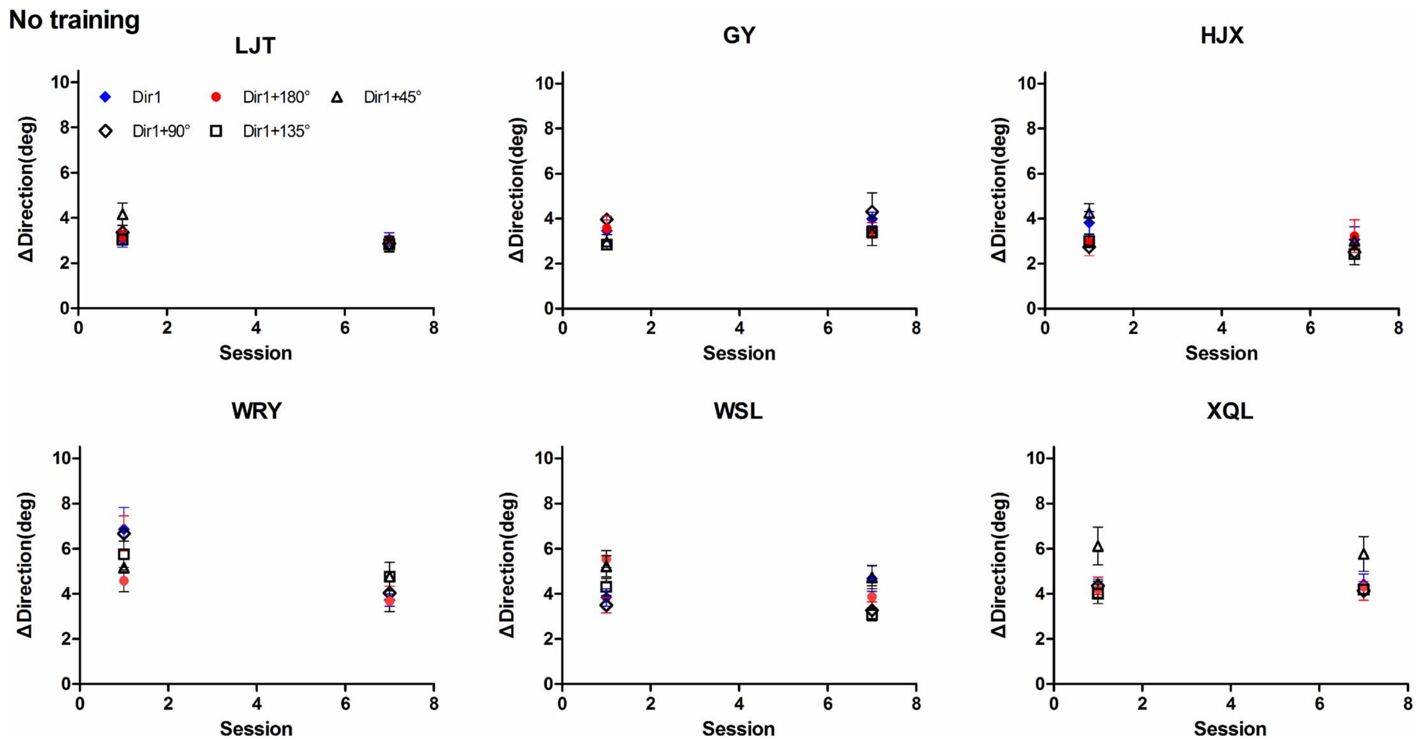

Figure 7. Data from six participants who went through no training but did only the pre- and posttraining measurement as the double- and single-training participants. These data were simultaneously collected with the last six participants' data in Figures 5 and 6, respectively, per request by a reviewer to avoid any "cohort effect." There was no reliable difference between the two sessions' thresholds.

Appendix B: The robust Bayesian *t* test on transfer indices

Measured in the transfer index (TI), the baseline transfer from our 30 participants in single training was 54% and that from the six participants in Zhang and Yang (2014, experiment 1) was 17%. The pooled baseline measure from these 36 participants was therefore 48%. The pooled double training TI from the current study ($N = 13$), Zhang and Yang (2014, experiment 2; $N = 6$), and Liang et al. (2015b; $N = 9$) was 74%. Was this TI of 74% in double training significantly greater than the TI of 48% in single training?

We asked this question because TIs, rather than discrimination thresholds, were exclusively used to address this question in the literature (Xiong et al., 2016; Zhang & Yang, 2014; Zhang & Yu, 2016). Because TI is defined as a ratio of ratios, its variability could be large. For example, LJC in double training in the current study had a TI as high as 251%. As far as we know, a transfer rate this high had not been established as real in the literature. Another problem of the TI was that any participant's pretraining threshold was treated effectively the same, which was a very strong assumption. For this reason, we preferred to use discrimination thresholds and *F* tests rather than TIs and *t* tests. We considered *t* tests here only for comparison purpose with prior studies.

In the TI measure, ZFY (182%) and ZCY (270%) in Zhang and Yang (2014) and XJJ (–66%) in Liang et al. (2015b) could also be considered as "outliers" (B. Klein, personal communications, September 7, 2016). These extreme values inflated the variability of transfer between pre- and posttraining. Rather than removing these outliers, we ran a robust Bayesian *t* test on this specific interaction contrast that accounted for these outliers by estimating the additional normality parameter for a *t*-distribution likelihood, thus providing a better estimate of transfer variance for each group of participants both in single and double training (Kruschke, 2013). We obtained a median difference of 24%. However, the highest density interval (95% HDI) of the posterior difference included zero (–5%, 56%), meaning there was little reliable difference in transfer between the single and double training groups. These results further confirmed our *F*-test results. Namely, no reliable benefit of double training could be found as compared with single training.

Given the data, the posterior distribution of the difference between conditions reflects the uncertainty of the single and double transfer effects. A larger amount of data (evidence) may overpower the vague priors, leading to greater precision of the median difference. To examine this possibility, we ran a Bayesian power analysis for a range of sample sizes to find the minimum number of participants needed to obtain a power of at least 0.8. The analysis suggests that at least 300 (power = 0.81 [0.78, 0.85]) but no more than 320 (power = 0.84 [0.81, 0.87]) participants would be needed in each condition to obtain a median difference of 0.25 between double and single conditions.

Appendix C: Discussion of Hung and Seitz (2014)

Per reviewer request, we analyze here Hung and Seitz (2014) that was set out to verify Zhang et al. (2010). The first task was Vernier discrimination, and the second task was orientation discrimination at the same retinal location. The question was, after the second task, whether Vernier discrimination could transfer to a new retinal location.

In their experiment 1, Hung and Seitz (2014) replicated the complete transfer result of Zhang et al. (2010). They then in experiment 2 changed both the first task's stimuli from Vernier Gabors to three-dot hyperacuity and the experimental method from staircase to constant stimuli. After these changes, the double training no longer enabled transfer. The authors suggested that this result was due to the change of the experimental method, rather than the change of stimuli. Nevertheless, in their experiment 3, the stimuli were changed back to Gabors as in experiment 1. This time, however, a single long staircase for the entire training was used that replaced the original multiple short staircases. The difference is that at the beginning of each short staircase, the Vernier offset was reset to a larger value. In the single long staircase, the offset was never reset and was at threshold (corresponding to 79.4% correct) for the large number of trials after initial learning. Hung and Seitz (2014) concluded that the no-transfer result was due to the large number of difficult trials at threshold in training.

Our analysis is as follows. Let us assume that the no transfer in experiment 2 was due to the change of method, as Hung and Seitz suggested, and not due to the change of stimuli. If the no transfer in experiment 3 was indeed due to the large number of difficult trials, then it is unclear why in experiment 2 there was no transfer either. In experiment 2, with the method of constant stimuli, five offsets were used: 0.9, 1.8, 2.7, 3.6, and 4.5 arcmin. The threshold at the end of training was about 2.5 arcmin (Figure 2B), which means that at least 20% of the trials during training were easy trials (offset = 4.5 arcmin). It is unclear why these easy trials were insufficient to enable transfer. In addition, as introduced earlier, Xiong et al. (2016) could maintain the difficulty level of a task (angular difference of motion directions) but jittered the directions within ±2°, to enable transfer. This indicates that the constant difficulty level in the single long staircase might not be responsible for the specificity.

There also appeared to be large data variation in Hung and Seitz (2014). In experiment 1 ($N = 6$), the pre- and posttraining thresholds were 10 and 7 arcmin, respectively (Figure 1B). However, in experiment 3 ($N = 6$), the pretraining threshold was already 7 arcmin (Figure 3B). In other words, the six participants in experiment 1 could reach only 7 arcmin as a result of a training method with many easy trials. In comparison, the six participants in experiment 3 had already achieved the 7 arcmin without any training. Finally, learning appeared to have occurred primarily from the pretraining measurement to the first training session (Figure 3B). Given that the threshold from the first training session was calculated from only the first 100 trials at most, it is unclear how much this learning was confounded by task familiarization. Therefore, it is unclear whether the lack of transfer in experiment 3 was due to a large number of difficult trials or due to insufficient learning.