Predictors of Sensitivity to Perceptual Learning in Children With Infantile Nystagmus

Bianca Huurneman,[1,2] F. Nienke Boonstra,[1,2] and Jeroen Goossens[1]

[1]Radboud University Medical Centre, Donders Institute for Brain, Cognition and Behaviour, Cognitive Neuroscience Department, Nijmegen, The Netherlands
[2]Bartiméus, Institute for the Visually Impaired, Zeist, The Netherlands

Correspondence: Bianca Huurneman, Cognitive Neuroscience Department, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Centre, Kapittelweg 29, 6525 EN Nijmegen, The Netherlands; B.Huurneman@donders.ru.nl.

Submitted: March 22, 2017
Accepted: July 26, 2017

Citation: Huurneman B, Boonstra FN, Goossens J. Predictors of sensitivity to perceptual learning in children with infantile nystagmus. *Invest Ophthalmol Vis Sci*. 2017;58:4162–4172. DOI: 10.1167/iovs.17-21913

Purpose. To identify predictors of sensitivity to perceptual learning on a computerized, near-threshold letter discrimination task in children with infantile nystagmus (idiopathic IN: $n = 18$; oculocutaneous albinism accompanied by IN: $n = 18$).

Methods. Children were divided into two age-, acuity-, and diagnosis-matched training groups: a crowded ($n = 18$) and an uncrowded training group ($n = 18$). Training consisted of 10 sessions spread out over 5 weeks (grand total of 3500 trials). Baseline performance, age, diagnosis, training condition, and perceived pleasantness of training (training joy) were entered as linear regression predictors of training-induced changes on a single- and a crowded-letter task.

Results. An impressive 57% of the variability in improvements of single-letter visual acuity was explained by age, training condition, and training joy. Being older and training with uncrowded letters were associated with larger single-letter visual acuity improvements. More training joy was associated with a larger gain from the uncrowded training and a smaller gain from the crowded training. Fifty-six percent of the variability in crowded-letter task improvements was explained by baseline performance, age, diagnosis, and training condition. After regressing out the variability induced by training condition, baseline performance, and age, perceptual learning proved more effective for children with idiopathic IN than for children with albinism accompanied by IN. Training gains increased with poorer baseline performance in idiopaths, but not in children with albinism accompanied by IN.

Conclusions. Age and baseline performance, but not training joy, are important prognostic factors for the effect of perceptual learning in children with IN. However, their predictive value for achieving improvements in single-letter acuity and crowded letter acuity, respectively, differs between diagnostic subgroups and training condition. These findings may help with personalized treatment of individuals likely to benefit from perceptual learning.

Keywords: perceptual learning, children's vision, low vision, congenital nystagmus

Visual perceptual learning, which refers to long-term improvements in visual abilities after training, has been regarded as a promising treatment option for a variety of vision defects, ranging from presbyopia[1] to hemianopia.[2,3] There is accumulating evidence that perceptual learning[4–7] might also be a promising treatment strategy to improve vision in subjects with low vision. One puzzling finding that could potentially hold back perceptual learning from entering clinical practice is the large interindividual variability in learning outcome.[8] For example, stereoacuity improvements in adults can range from no improvement at all to a 30-fold improvement, with the median learning effect lying around a 1.7.[9] In a recent study on visual perceptual learning in children with infantile nystagmus (IN), we also found large variability in perceptual learning outcomes.[5–7] Some children did not show any significant changes in visual performance while others showed an almost 8-fold improvement. Why is it that some children benefit so much while others do not?

Recent studies have succeeded in explaining some of the interindividual variability in the magnitude of learning effects. Factors that have been identified to explain the magnitude of visual perceptual learning include baseline performance (poor initial performers tend to show larger improvements, typically explaining 40% to 60% of the outcome variance),[8–12] number of training sessions,[11,13] age[14] (but note that there are also studies that found no effect of age[15–17]), motivation and intelligence,[18] performance feedback,[19–21] the accuracy of performance during training,[22] heart-rate variability,[23] and reactivity of dopaminergic and cholinergic neuromodulatory systems.[24,25] The goal of the present study was to evaluate the contribution of physiological factors, such as baseline performance, age, and clinical diagnosis, and a psychological factor, the perceived pleasantness of the training (training joy). Toward this end we evaluated to what extent the interindividual variability in perceptual learning that we observed previously[7] might be accounted for by baseline performance, diagnosis, age, training condition, and training joy of the participants.

Our visual perceptual learning paradigm implemented several features to evoke more generalization of learning and enhance motivation during training.[7] To evoke more generalized learning, the training task combined two approaches: (1)

identifying a letter at the acuity threshold and (2) execution of a goal-directed saccade toward the target letter under time constraints. To enhance the children's motivation we included trial-by-trial feedback about their performance by presenting happy- and neutral-face smileys for correct and incorrect responses, respectively, and a "Get the picture" game that they could play between training blocks as a reward. Children performed the training under either crowded or uncrowded conditions. Both training groups showed improved crowded and uncrowded visual acuity (VA) after ~3 1/2 hours of training. Mean improvements on different VA tasks ranged between 0.085 and 0.15 logMAR, which is in line with studies in other patient groups using the same number of training hours.[10,26] Moreover, the training benefits from both the crowded and uncrowded training generalized to significant improvements in stereopsis[7] and spatial aspects of reading.[5] In fact, the improvements in reading acuity, critical print size, and acuity reserve were tightly linked to the improvements in uncrowded distance acuity and the reduction in crowding extent.[5] By contrast, fixation stability, saccade behavior, and nystagmus properties were unaltered by the training, suggesting that the vision improvements were primarily due to improved sensory processing rather than improved oculomotor behavior.[6] The key question for the present study is therefore which other variables might predict the training-induced improvements in sensory processing in children with IN.

Perceptual learning seems to be gated by reinforcement processes,[27,28] where internal motivation can work as an important reinforcer and successful performance of a task acts as an internal reward.[29,30] The dopamine system is thought to mediate the neuropharmacology of reward[31]; if we succeed in accomplishing a task ("Yes, I got the correct answer!"), this will trigger mesolimbic dopamine neurons to release dopamine, resulting in a natural high.[32] Selective attention, that is, the process of selecting and prioritizing a small subset of sensory information for processing,[33] plays an important role in perceptual learning.[34,35] Perceptual learning is more pronounced for features that are attended or cued than for unattended features.[33] The attention-gated reinforcement learning (AGREL) model[24] proposes that two factors guide learning: (1) a global neuromodulatory signal that informs all synapses whether the outcome of a trial is better or worse than expected, and (2) an attentional feedback signal from the response selection stage, which restricts plasticity to those areas that were responsible for the behavioral choice. This suggests that perceptual learning might be more effective if subjects are enjoying the task more and pay more attention to it.

In addition to assessing the impact of baseline performance, age, diagnosis, and training condition on the magnitude of learning, we therefore evaluated the notion that training joy ("I like this") might matter for perceptual learning performance. Learning outcome was quantified as the before–after difference in visual discrimination performance on the crowded- and uncrowded-letter discrimination task. We hypothesized that baseline performance, in line with previous findings in adults,[8–10] is a significant predictor of perceptual learning in children with IN and that more training joy predicts a better training outcome.

METHODS

Participants

The present study was part of a larger project in which we assessed the effects of visual perceptual learning in children with IN on visual performance,[7] reading performance,[5] and oculomotor behavior.[6] We used the data from the same 36 children with IN that were included in those previous studies (idiopathic IN [IIN]: $n = 18$; oculocutaneous albinism accompanied by nystagmus [ALB/IN]: $n = 18$). Inclusion criteria were age 6 to 11 years old, normal birth weight, birth at term, no perinatal complications, normal development, and no motor or mental impairments. Children with visual impairment and a distance VA poorer than 1.3 logMAR or better than 0.3 logMAR were excluded. Participants in the crowded and uncrowded training groups were age and acuity matched (crowded training group: mean age ± SD: 111 ± 20 months, mean acuity ± SD: 0.54 ± 0.28 logMAR; uncrowded training group: mean age ± SD: 109 ± 18 months, mean acuity ± SD: 0.58 ± 0.25 logMAR). The parents of all children gave informed consent after explanation of the purpose, procedure, and possible consequences of the study. The study was reviewed and approved by the local ethics committee (CMO Arnhem-Nijmegen) before the measurements and training started. The study adhered to the principles of the Declaration of Helsinki.

Ophthalmologic Exam

A short ophthalmologic assessment was performed at pretest and post test for all children (see Ref. 7 for methodological details). Monocular and binocular distance VA was measured with the Landolt C-test at 5 m. The C-test had an uncrowded chart, with an interelement spacing of at least 30′, and a crowded chart, with an interelement spacing of 2.6′.[36] Binocular near VA was measured at 40 cm with the LH-version of the C-test.[37]

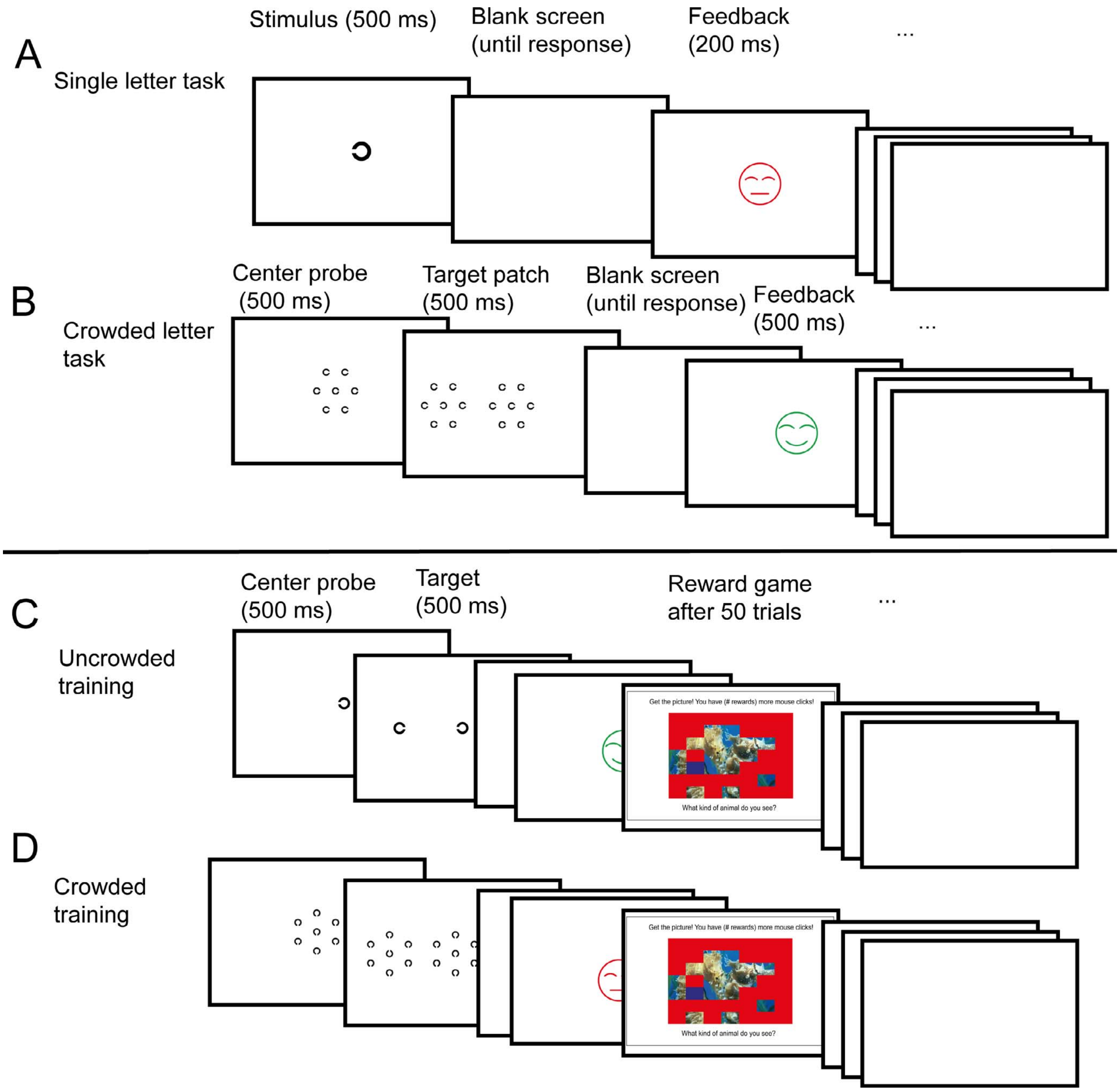
Experimental Tasks: Single- and Crowded-Letter Task

In addition to the ophthalmologic measures, every child performed a computerized single-letter and crowded-letter discrimination task. Stimuli in these tasks were high-contrast black Landolt Cs (0.3 cd/m^2) with four possible orientations on a white background (193.8 cd/m^2).

Single-letter VA was measured with the single-letter task (Fig. 1A) in which the children had to indicate the orientation of isolated Landolt C rings that were presented for 500 ms at the center of the screen. The computer used the uncrowded distance VA measured during the ophthalmologic examination as a starting point and randomly showed Landolt Cs in six different sizes ranging from 0.2 log units below to 0.3 log units above the uncrowded distance VA. Each size was presented 15 times so there were a total of 90 (15 × 6) trials. From the resulting scores, the 62.5% correct uncrowded-letter discrimination threshold was determined (see Ref. 7 for details).

The crowding extent was measured with the crowded-letter task (Fig. 1B). The crowding extent refers to the spacing between target and flanker that a child needed to identify the target in the middle of a group of seven Landolt C rings that appeared at the left or right side of the screen (Fig. 1B). The target was always presented in a different orientation than the distractors (which had the same orientation). Letter size was kept fixed at 0.15 log units above the subject's uncrowded acuity threshold. Center-to-center spacing of the target and flankers was varied to determine the 62.5% correct discrimination threshold (see Ref. 7 for details).

During the single- and crowded-letter tasks, children were seated at 130 cm from the screen if their uncrowded distance VA was $\leq$0.70 logMAR, and the distance between the central and target stimulus was 5° to prevent the second patch from falling outside the screen. Children with uncrowded distance



Get the picture! You have (# rewards) more mouse clicks!

What kind of animal do you see?

Get the picture! You have (# rewards) more mouse clicks!

What kind of animal do you see?

FIGURE 1. Experimental tasks and training paradigms. (**A**, **B**) Impression of the single-letter task (**A**) and the crowded-letter task (**B**) that were performed at pre- and posttest. Children were asked to indicate the orientation of the (central) Landolt C. Letter size or letter spacing was varied to determine discrimination thresholds. (**C**, **D**) Impression of the uncrowded training task (**C**) and crowded training task (**D**). As in the crowded-letter task, subjects initially looked at a central probe. After 500 ms, the target stimulus was briefly (500 ms) presented either to the left or to the right of the central probe, and children had to identify the orientation of the Landolt C in the center of this peripheral patch. Letter size or spacing and letter size were gradually reduced based on a child's improvements in performance during the training. Feedback was given after each trial. In between training blocks, children played a reward game.

VA > 0.70 logMAR were seated at 50 cm and the distance between the central and target stimulus was 15° to prevent the patches from overlapping.

Training

Children were randomly divided into two age-, acuity-, and diagnosis-matched training groups: a crowded training group and an uncrowded training group. Children in the crowded training group trained with the same crowded-letter task that was performed before and after training, but during training the target-to-flanker spacing and letter size were adjusted based on the child's performance (Fig. 1D). In the uncrowded training group, the training routine was very similar to the crowded training routine, except that in the uncrowded training condition, Landolt C rings were presented in isolation instead of grouped (Fig. 1C). In the uncrowded training group, letter size was the only variable that was adapted based on performance improvements. The rationale behind implementing the two-step routine of first presenting a central patch followed by the presentation of the target patch on the left or right was to train oculomotor and attention control and possibly enhance generalized learning by also training these two skills. Average reaction times of the subjects' saccadic orienting responses toward the target stimulus were around 400 ms.[6]

After each block of 50 trials, the children could enjoy a mini-game called "Get the picture." After the first block, the children were shown a picture masked by a 7 × 7 rectangle matrix covering the whole picture (see Figs. 1C, 1D). The child was allowed to click the mouse one to seven times after each block, revealing the picture piece by piece. The number of clicks depended on the performance: A better performance meant more clicks. After the last block, the child could click as often as needed to reveal the whole picture. Furthermore, trial-by-trial feedback was provided by means of a happy green smiley or a neutral red smiley for a correct or incorrect answer, respectively. One training session consisted of 350 trials. Children were trained twice a week for 5 weeks, leading to a total of 3500 trials per training. See Huurneman et al.[7] for further details about the training protocol.

Training Joy

At the end of each training session, children rated how much they enjoyed or did not enjoy that particular session by pointing out a smiley on a 5-point smiley scale. The child was asked the question "How much fun was today's training?" The 5-point score is based on the Smiley Scale, which has been developed especially for use with younger children by the Junior School Project team to measure children's attitudes toward different aspects of school life and the curriculum.[38] It is easily administered and enables attitude to be assessed even when children are not fluent readers. The scale uses happy to sad faces to represent each point on the scale, which ranges from very negative (score of 1) to very positive (score of 5).[38] The Smiley Scale is a reliable instrument for the testing of young children's attitude.[39]

Procedure

Training began within 2 weeks after the pretests. In total, all children trained twice a week for 5 consecutive weeks, for a total of 10 training sessions. All 20-minute training sessions were performed at the child's school. Within 2 weeks after the last training session, posttests took place.

Apparatus

The experimental tasks were run on a Dell M4700 15.6-inch laptop (Round Rock, TX, USA) and written in a Matlab environment (MathWorks, Inc., Natick, MA, USA) using the Psychophysics Toolbox version 3.0.12 extensions.[40] During pre- and posttest measurements, stimuli were displayed on a 32-inch Dell Ultra HD monitor (UP3214Qt 32″; 3840 × 2160 pixels; pitch 0.18 mm^2). The training was executed on a 15.6-inch Dell M3800 laptop (3200 × 1800 pixels; pitch 0.11 mm^2).

Statistical Analysis

Statistical analysis was done in Matlab (version 2014b; MathWorks, Inc.) using the Statistical Toolbox. We focused our analyses on the computerized uncrowded distance VA and crowding extent measures rather than on the vision chart measures for three reasons: (1) The computerized measures were more objective as they could not be influenced by the experimenter; (2) they were more sensitive and robust because the step sizes were smaller and the number of trials to estimate the thresholds was larger; and (3) the single-letter and crowded-letter computer tasks were closest to the trained task. Our oculomotor study showed that training induced improvements in visual performance are decoupled from the subjects' eye movement behaviour[6] and therefore are not considered as predictors for perceptual learning in the present study. Our first step was to conduct a 3-way analysis of variance with training condition and diagnosis as independent variables to evaluate whether there were main differences and/or interactions between training groups and/or diagnostic groups with regard to our two outcome measures, training-induced changes in single-letter VA and training-induced changes in crowding extent.

The next step was to conduct multiple linear regression analyses to model the training effects as a function of baseline performance (VA or crowding extent in logMAR), age (in months), training joy (mean smiley score across all 10 training sessions), diagnosis (ALB/IN [coded as 0], idiopathic IN [coded as 1]), and training condition (uncrowded [coded as 0] and crowded training group [1]). Note that the latter two variables are dichotomous variables (diagnosis and training group) whereas the first three are continuous variables (baseline performance, age, and training joy). Baseline performance, age, and training joy were not significantly correlated with one another (Pearson's correlations $r < 0.41$). Missing values for training joy ($n = 4$) were imputed by estimating their values from the average of two imputation methods: regression imputation and mean substitution.

We used the total sum of squared error (SSE) to select variables for the best regression model. This method applied an *F*-test to evaluate if adding or removing terms yielded a significantly better or worse model fit given the different number of free parameters it included. Our starting model contained an intercept, linear terms for each predictor and all products of pairs of those predictors (no higher-order interaction terms). If the P value for the *F*-test of the change in the SSE caused by adding a term was smaller than 0.01, then the term was added. If the P value of the *F*-test for adding a term was larger than 0.05, the term was removed from the model. We used these relatively strict variable selection criteria (more common values for variable inclusion and exclusion are $P < 0.05$ and $P > 0.10$, respectively) to prevent overfitting our limited datasets. Values of the continuous variables were centered on their respective means. In this way, the regression coefficients associated with the two dichotomous grouping variables reflect the effect of training condition and diagnosis on an average child (average in terms of baseline performance, age, and training joy across all included children). The intercept represents the improvement induced by the uncrowded training in an average child with ALB/IN.

The contribution of a variable to the overall variance was expressed by the partial correlation coefficient or the percentage of the SSE that was uniquely explained by this variable. Contributions were considered statistically significant if the P value of the corresponding *t*-test was <0.05. Note that the standard errors of the regression coefficients, and hence the *t*-statistics, also incorporate the inevitable within-subject test–retest variability because this variability adds to the overall variance of the (within-subject) before–after difference measures.

In case one of the continuous variables showed a significant interaction with a dichotomous variable, this was considered an indication that the continuous variable had a different effect in the two groups. To determine its influence in each group, analyses were stratified by group.

RESULTS

On average, the crowded and uncrowded training paradigms resulted in very similar, positive training outcomes[7]: 10 training

Table 1. Mean Improvements in VA and Crowding Extent (Mean ± Standard Error of the Mean) in Children With Albinism and Children With Idiopathic IN (IIN) After Crowded and Uncrowded Training, Respectively

	Albinism		IIN		Condition and/ or Diagnosis Effects
	Crowded	Uncrowded	Crowded	Uncrowded	
N	6	8	10	8	
Δ VA, logMAR	0.057 ± 0.021	0.099 ± 0.027	0.060 ± 0.029	0.118 ± 0.030	n.s.
N	8	10	10	8	
Δ CE, logMAR	0.252 ± 0.086	0.144 ± 0.055	0.344 ± 0.067	0.260 ± 0.077	n.s.

CE, crowding extent; n.s., not significant.

sessions (equal to ~3.3 hours of training) resulted in a significant improvement in both single-letter VA (0.085 ± 0.014 logMAR) and crowding extent (0.25 ± 0.04 logMAR). As shown in Table 1, the average improvements in VA and crowding extent were statistically significant for all subgroups (single-letter VA: $F(1,28) = 31.68$, $P < 0.001$; crowding extent: $F(1,30) = 46.13$, $P < 0.001$), but this group-level analysis revealed no significant differences between the four groups, that is, there were no interactions between or main effects of training condition and diagnosis (all $P > 0.05$). There were four missing values for the changes in single-letter VA. Two of these children had ALB/IN and completed the crowded training (ID 1 and 2), and two of these children had ALB/IN and completed the uncrowded training (ID 5 and 7). There were also two missing values for changes in crowding extent (ID 5 and 7).

While progress at the group level was very similar, the magnitude of the improvement varied considerably between subjects. To illustrate this variability across subjects, Figures 2

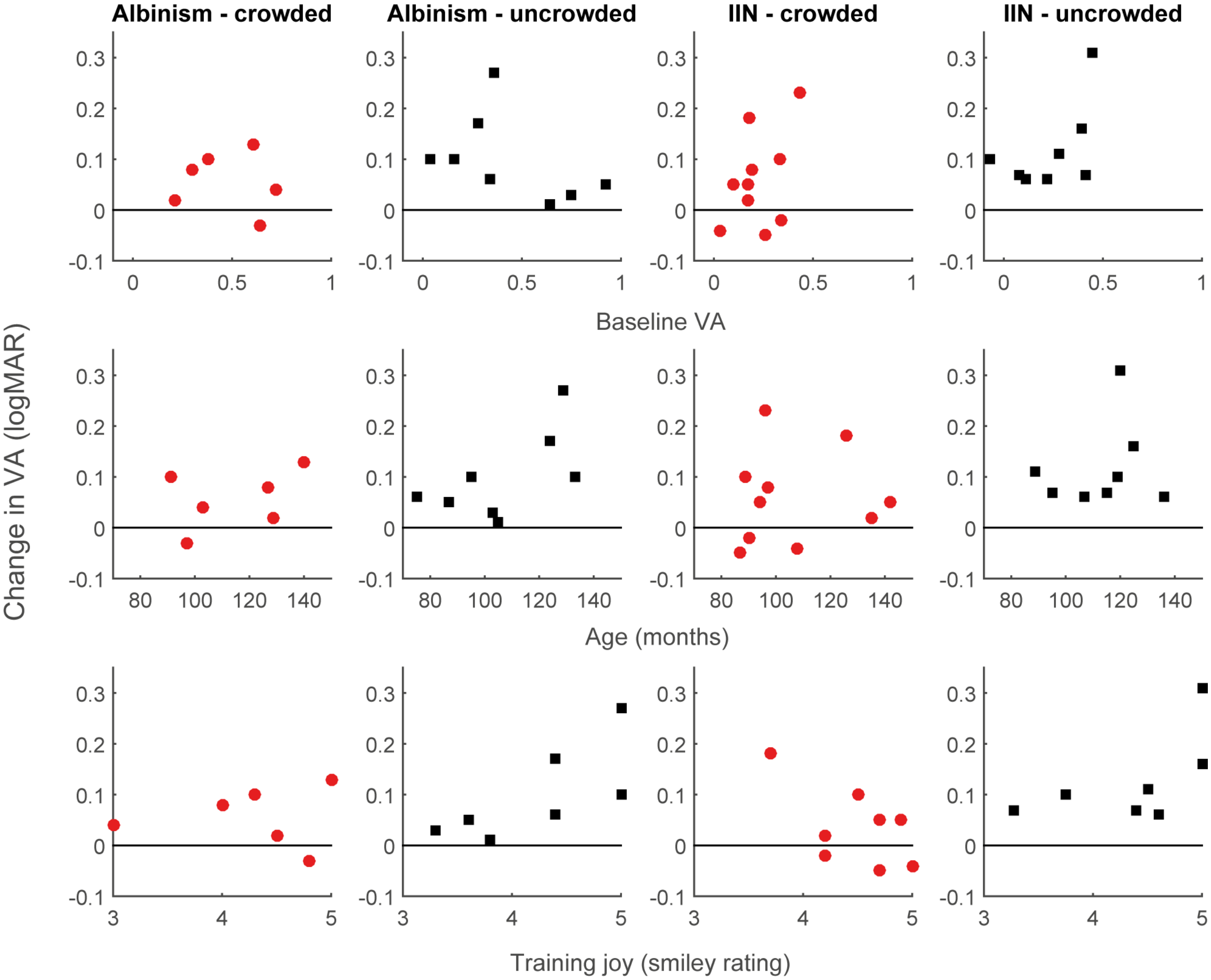


Figure 2. Scatter plots of training-induced changes in visual acuity (VA) as a function of baseline visual acuity, age, and training joy. Positive values indicate improved performance after training. Data are stratified by diagnoses (albinism and idiopathic IN) and training condition (crowded and uncrowded).

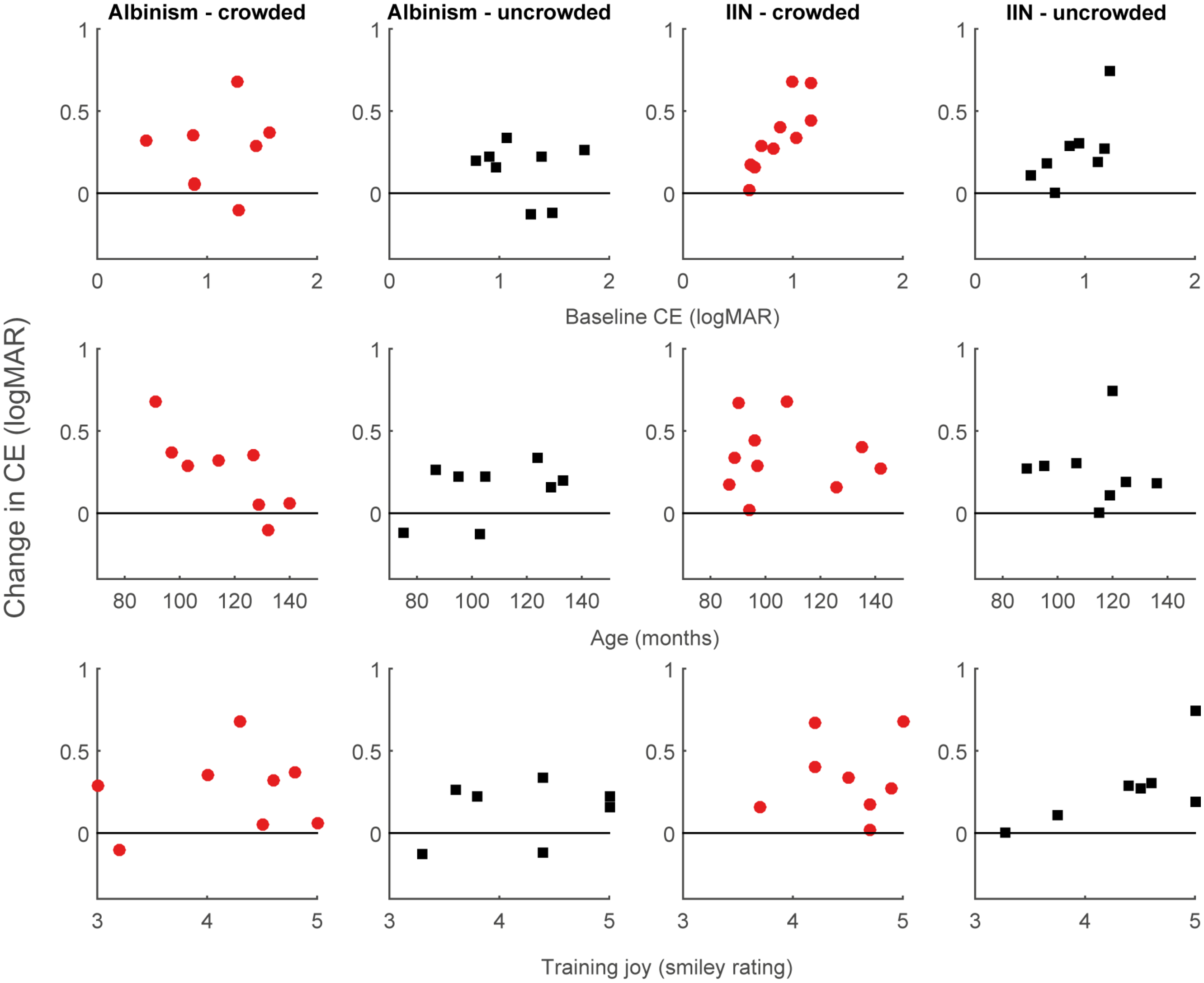


FIGURE 3. Scatter plots of training-induced changes in crowding extent (CE) as a function of baseline visual acuity, age, and training joy. Positive values indicate improved performance after training. Data are stratified by diagnoses (albinism and idiopathic IN) and training condition (crowded and uncrowded).

and 3 show scatter plots of the change in VA and crowding extent as a function of the three continuous predictor variables that were considered in our analyses. Data are stratified by the two dichotomous variables, training condition and diagnosis. Our next step was to model this interindividual variability by assessing the contributions of physiological and psychological predictors on the training outcomes using multiple linear regression analyses.

Although single-letter acuity and crowding extent were significantly correlated (before training: $r = 0.80$, $P < 0.001$; after training: $r = 0.75$, $P < 0.001$), the training-induced improvements in these two measures were not ($r = 0.10$, $P = 0.581$). In line with this observation, we found that the intersubject variability in these two measures of training benefit was best described by two distinct regression models.

Training Outcome: Predictors of Changes in Single-Letter Acuity

In the best-fitting regression model for improvements in single-letter acuity, 57% of the total variance could be explained by age, training joy, and training condition ($F(4, 27) = 8.86$, $P < 0.001$; Table 2; Fig. 4). Older age was associated with larger single-letter acuity improvements ($\beta = 0.018 \pm 0.007$ logMAR/year, $t(27) = 2.60$, $P = 0.015$, partial correlation $r = 0.45$, $R^2 = 0.20$). Training joy, on the other hand, had a different effect in the two training conditions (training joy × training condition interaction: $t(27) = -4.96$, $P < 0.001$). A post hoc analysis was therefore conducted to further characterize the effect of training joy in the two training conditions.

In the uncrowded training group, children who enjoyed the training more showed larger single-letter acuity improvements ($\beta = 0.09 \pm 0.02$ logMAR/point, $t(13) = 3.98$, $P = 0.002$, partial correlation $r = 0.74$, $R^2 = 0.55$). In the crowded training group, however, improvements in single-letter VA decreased with increasing training joy ($\beta = -0.06 \pm 0.02$ logMAR/point, $t(13) = -2.98$, $P = 0.011$, partial correlation $r = -0.64$, $R^2 = 0.41$). Training condition itself explained 17% of the variability; after correcting for the effects of training joy and age, children in the uncrowded training group showed larger single-letter acuity improvements than children in the crowded training group (difference: $\beta = -0.05 \pm 0.02$ logMAR, $t(27) = -2.33$, $P = 0.027$).

Table 2. Parameters of the Linear Regression Model That Predicted Changes in Visual Acuity (VA in logMAR) Best

Predictor	β (SE)	*t*-Test	*P* Value	Partial Correlation	R^2	R^2_{adj} *
Intercept	0.109 (0.015) logMAR	7.41	<0.001	n.a.	0.57	0.50
Age	0.017 (0.007) logMAR/y	2.60	0.015	0.45*		
Training joy	0.09 (0.02) logMAR/point	3.65	0.001	0.57†		
Training condition	−0.05 (0.02) logMAR	−2.33	0.027	−0.41*		
Training joy × training condition	−0.16 (0.03) logMAR/point	−4.96	<0.001	−0.69‡		

Age and training joy were centered on their means ([mean ± SD] 9.2 ± 1.5 years, and 4.23 ± 0.69 points, respectively). adj, R^2 adjusted for the number of predictor variables in the model; n.a., not applicable.

* 2-tailed $P < 0.05$.

† 2-tailed $P < 0.01$.

‡ 2-tailed $P < 0.001$.

Training Outcome: Predictors of Changes in Crowding Extent

The best-fitting model for changes in crowded extent, which included baseline crowding extent, age, diagnosis, and training condition as predictors, accounted for 56% of the total variance ($F(7, 26) = 4.70$, $P = 0.002$; Table 3; Fig. 5). The effect of baseline performance and age was different in the two diagnostic groups (i.e., baseline crowding extent × diagnosis interaction: $t(26) = 3.78$, $P < 0.001$; age × diagnosis interaction: $t(26) = 2.11$, $P = 0.044$). In addition, age had a different effect in the two training conditions (age × training condition interaction: $t(26) = -2.74$, $P = 0.011$). To investigate these interaction effects, post hoc analyses were run.

We first split the data by diagnosis to evaluate the diagnosis-dependent effects of baseline crowding extent and age. In the group of children with ALB/IN, age affected the outcome of the two types of training in a different manner (age × training condition interaction: $\beta = -0.18 \pm 0.05$ logMAR/year, $t(11) = -3.30$, $P = 0.007$, partial correlation $r = -0.70$). The uncrowded training improved the crowding extent of children with ALB/IN (intercept: 0.17 ± 0.06 logMAR), but none of the predictors could explain any interindividual differences in its outcome for this group ($P > 0.1$). The effect of the crowed training, on the other hand, was on average larger ($\beta = 0.35 \pm 0.05$ logMAR, $t(5) = 6.47$, $P = 0.001$) and significantly correlated with age ($\beta = -0.16 \pm 0.04$ logMAR/year, $t(5) = -4.14$, $P = 0.009$); younger children with ALB/IN in the crowded training group showed larger improvements. In the idiopathic IN group, crowding extent improved equally under the two training conditions (intercept: 0.34 ± 0.06 logMAR), and poorer baseline performance was associated with larger improvements ($\beta = 0.71 \pm 0.17$, $t(13) = 4.28$, $P < 0.001$, partial correlation $r = 0.76$). The other terms did not correlate with the interindividual changes ($P > 0.1$).

In the second post hoc analysis we split the data by training condition to evaluate the different effects of age on the uncrowded and crowded training. Age did not predict the changes in crowded extent for children in the uncrowded training group ($P > 0.3$). By contrast, age did matter in the crowded training group: Younger children showed larger improvements ($\beta = -0.16 \pm 0.04$ logMAR/year, $t(13) = -4.22$, $P = 0.001$, partial correlation $r = -0.77$). However, the analysis above (with data split up by diagnosis) demonstrates that age influenced the effect of the crowded training only in children with ALB/IN.

After eliminating the variability that could be explained by age and baseline performance, it appeared that idiopathic IN gave a better outcome prognosis compared with ALB/IN ($\beta = 0.16 \pm 0.07$ logMAR, $t(26) = 2.38$, $P = 0.025$). Likewise, we found that the crowded training was more effective in improving the crowding extent than the uncrowded training ($\beta = 0.14 \pm 0.06$ logMAR, $t(26) = 2.45$, $P = 0.021$).

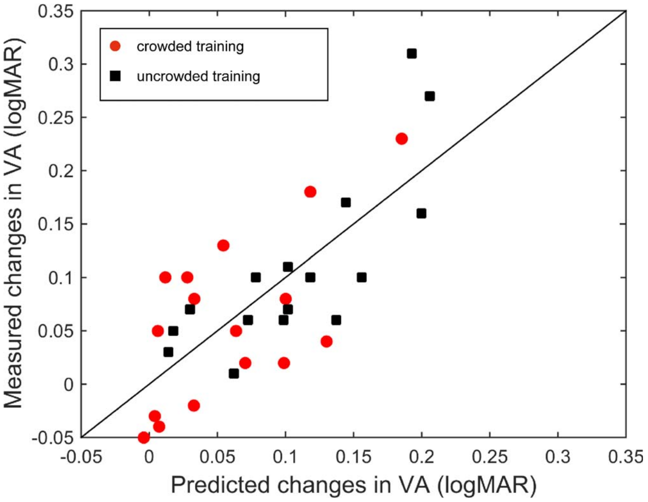


Figure 4. Scatter plot of measured versus predicted changes in visual acuity. Our linear regression model (Table 2) predicted 57% of the total variance in changes in single-letter visual acuity (VA). *Red circles* correspond with the crowded training condition and *black squares* with the uncrowded training condition. The *solid line* is the identity line.

DISCUSSION

In the current study we tested if baseline performance, age, diagnosis, training condition, and training joy can predict the sensitivity to perceptual learning in children with IN. In line with recent findings in patients with amblyopia,[11,13] our results show that there are indeed measurable factors that can predict the amount of improvement that children with IN have from perceptual learning. We hypothesized that lower baseline performance and more training joy would predict better training outcomes. Our results provide partial support for these hypotheses, but the relationships proved to be more complex, showing different predictors for different outcome measures of the training and different effects of the predictors for different subsets of subjects and different types of training. The difference in predictor subsets for the two outcome measures might be explained by the longer maturational time window of crowded- compared to single-letter VA and the higher sensitivity of crowded- versus single-letter tasks to pick up inaccurate attentional and gaze control.[37,41,42] Below we elaborate on the predictive value of the different factors that we studied.

TABLE 3. Parameters of the Best Linear Regression Model for Changes in Crowding Extent (CE in logMAR)

Predictor	β (SE)	*t*-test	*P* Value	Partial Correlation	R^2	R^2_{adj}
Intercept	0.17 (0.05) logMAR	3.19	0.003	n.a.	0.56	0.44
Baseline CE	−0.13 (0.16)	−0.84	0.41	−0.16		
Age	−0.002 (0.038) logMAR/y	−0.05	0.96	−0.01		
Diagnosis	0.16 (0.07) logMAR	2.38	0.025	0.42*		
Training condition	0.14 (0.06) logMAR	2.45	0.021	0.43*		
Baseline CE × diagnosis	0.88 (0.23)	3.78	<0.001	0.60†		
Age × diagnosis	0.09 (0.04) logMAR/y	2.11	0.044	0.38*		
Age × training condition	−0.11 (0.04) logMAR/y	−2.74	0.011	−0.47*		

Baseline CE and age were centered on their means ([mean ± SD] 1.03 ± 0.33 logMAR, and 9.2 ± 1.5 years, respectively). adj, R^2 adjusted for the number of predictor variables in the model; CE, crowding extent; n.a., not applicable.

* 2-tailed $P < 0.05$.

† 2-tailed $P < 0.001$.

Baseline Performance

A tentative explanation for the finding that a larger crowding extent at baseline is predictive of a better training outcome only in children with idiopathic IN might be that children with ALB/IN typically have significant associated afferent sensory defects. Subjects with oculocutaneous albinism, the form of albinism that all children included in our study were diagnosed with, suffer in variable degrees of hypopigmentation and have ocular abnormalities such as iris transillumination, foveal hypoplasia, optic nerve head abnormalities, and abnormal, increased crossings of optic nerve fibers.[43] In individuals with idiopathic IN, there is not such an extensive retinal defect that limits vision.[44,45] Neuroimaging studies indicate that subjects with albinism also show reduced gray matter volume in the posterior occipital cortex (associated with the foveal representation),[46,47] a shorter calcarine fissure,[48] increased cortical thickness in the posterior part of V1 (negatively correlated to VA in albinism), and decreased gyrification in the left ventral occipital lobe compared to controls.[47] Cortical thickness in V1 is also increased in early-blind individuals, suggesting reduced pruning of synapses in V1 during the critical period.[49] A lot of these changes in brain morphology are considered to be the consequence of reduced visual input. In addition, functional magnetic resonance imaging studies suggest that the cerebellum is a possible site involved in the oculomotor dysfunction associated with idiopathic IN,[50,51] but there are no findings of severely altered brain morphology in subjects with idiopathic IN. Taken together, this suggests that perceptual learning in poor-sighted subjects with albinism might be limited by fundamental abnormalities of the visual system, whereas there are no such abnormalities described for subjects with idiopathic IN.

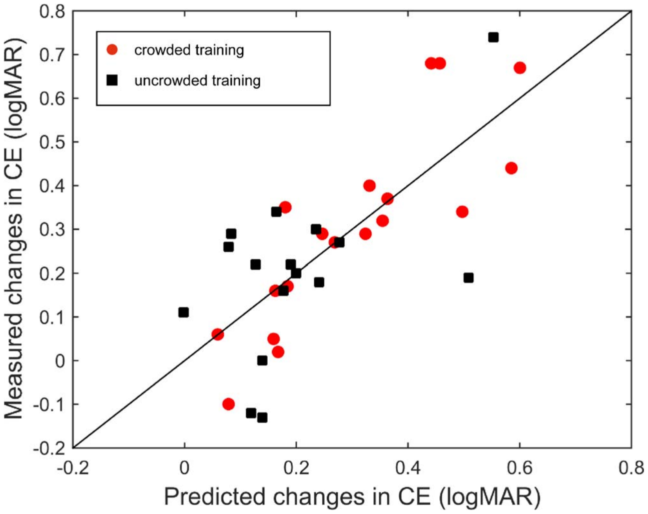


FIGURE 5. Scatter plot of measured versus predicted changes in crowding extent. Our linear regression model (Table 3) predicted 56% of the total variance in changes in crowding extent (CE). *Red circles* correspond with the crowded training condition and *back squares* with the uncrowded training condition. The *solid line* is the identity line.

The apparent absence of a relation between baseline single-letter VA and improvements on the single-letter task might be due to the simpler task demands of the single-letter task compared to the crowded-letter task. The presence of distracters and the oculomotor demand reduced performance by ~0.2 logMAR and made it a more challenging task where there was more room for improvement possible. However, it should be noted that a recent study[4] also did not find a significant link between pretraining performance and training-induced improvements in 9- to 18-year-old subjects with low vision.

Diagnosis

Our analysis revealed that after factoring out the differential effects of age and baseline, children with idiopathic IN did show significantly larger improvements on the crowded-letter task than children with ALB/IN. An explanation for this difference might be that perceptual learning is limited in children with albinism due to the presence of profound retinal/sensory deficits and alterations in the retinocortical visual pathway.

Age

Reviews on the effect of age on perceptual learning often report that age does not systematically account for variance in training outcome.[15,52] Here, the correlation between age and training effects were different for the two outcome measures: Age showed a strongly positive correlation with improvements in single-letter VA but a strongly negative correlation with improvements in crowding extent in the crowded training group.

A tentative explanation for the positive correlation between age and VA improvements in single-letter VA is that older children have stronger visual selective attention skills[53] that could boost learning. Reweighting theories postulate that perceptual learning involves the reweighting of connections between early- and mid-/high-level brain areas over the course of the training and that higher level visual cortical areas contribute to this process.[54] Furthermore, evidence is rising that perceptual learning operates via a reduction of internal neural noise and/or through more efficient use of the stimulus information by retuning the weighting of the information in higher-level visual decision areas.[52] These decision areas are not fully developed yet the age of 6 or 7 years.[55] The negative correlation between age and improvements on the crowded-letter task might be explained by the difficulty of the crowded training task. At baseline, young children struggled more with the crowded-letter task than the single-letter task. It is unclear why we observed this age effect only in the group of children with ALB/IN but not in the idiopathic IN group.

Training Condition

After controlling for the influence of other variables to the overall variability in training outcome, training condition did affect training outcome. More specifically, the uncrowded training proved to be more effective in improving the single-letter VA, while the crowded training was more effective in reducing the crowding extent. This finding could be expected, since one of the hallmarks of perceptual learning is the task specificity of learning effects.[56–59]

Training Joy

Previous studies have suggested a positive impact of training joy on training outcome, but, to our knowledge, this relation has never been evaluated quantitatively.[52,60] Our present results do not offer support for the idea that there is a general positive effect of training joy on training outcome. In fact, we found an opposite relation between self-reported training joy and single-letter VA improvements for the two training conditions: A positive relation was observed for the uncrowded training, while a negative relation was found for the crowded training. If experiencing more training joy reinforces the learning of specific, task-relevant features,[24] one would expect a positive relation between training joy and performance improvements on that trained task, but not necessarily between training joy and performance improvements on a (slightly) different, untrained tasked. However, this does not explain the negative association between training joy during the crowded training and improvements on the untrained, single-letter task. Moreover, training joy had no significant relation with the training-induced changes in crowding extent. We are confident that these findings are not an artifact of our imputation method to estimate the values of the four missing training joy scores; very similar results were obtained when we eliminated the four subjects with missing scores from our regression analyses. We should note, however, that there was limited interindividual variability in perceived joy; children generally rated the training paradigms as very enjoyable (average rating 4.3 ± 0.1 [SEM] on a scale of 1 to 5). Future research should adopt a study design in which the causal relation between training joy and progress can be evaluated.

Limitations

Our results indicate that training outcome cannot be predicted in a reliable manner by one single predictor; the relation between predictors and training outcome appears to be more complex. It is possible that the true relations are even more intricate, and that factors other than the ones included in our models do contribute. We cannot exclude, for example, that nystagmus frequency is also a predictor for the success of perceptual learning in children with IN. However, given the correlation between nystagmus frequency and diagnosis (nystagmus frequency was systematically lower in children with albinism than in children with idiopathic IN[6]), it is difficult to dissociate between these two factors on the basis of our current dataset. Indeed, a clear limitation of the present study is the relatively small number of subjects in each subgroup. Note, however, that we used stringent variable selection criteria to limit the complexity of our models and focus on the strongest predictors that could help the clinician to select patients for whom perceptual learning could be a successful treatment strategy. If we had used more liberal selection criteria, there would be a higher risk of overfitting our data. Overfitting can occur when a regression model is excessively complex relative to the amount of data available and threatens the validity of inferences and predictions made by the model.[61]

CONCLUSIONS

Our findings support the notion that poorer initial performance predicts larger improvements in children with idiopathic IN. Surprisingly, this effect of baseline performance on improvements in the crowded-letter task was not found in children with albinism accompanied by nystagmus. Furthermore, our results indicate that diagnosis matters: Children with idiopathic IN showed larger improvements than children with albinism accompanied by nystagmus on the crowded-letter task. Against our initial expectations, we found no support for the proposition that training joy systematically enhances training outcome. Although we used only a modest amount of training, our regression analyses did reveal the contributions of age, baseline performance, and diagnosis in explaining variability in perceptual learning outcome in children with IN. Our study is the first to identify diagnosis and baseline performance as possible prognostic factors for perceptual learning outcome in children with IN.

Acknowledgments

The authors thank the parents and children for their participation.

Supported by the ODAS Foundation and Landelijke Stiching Blinden en Slechtzienden (LSBS) which contributed through UitZicht, and Bartiméus Sonneheerdt. The authors alone are responsible for the content and writing of the paper.

Disclosure: **B. Huurneman**, None; **F.N. Boonstra**, None; **J. Goossens**, None

References

1. Polat U, Schor C, Tong JL, et al. Training the brain to overcome the effect of aging on the human eye. *Sci Rep*. 2012;2:278.
2. Das A, Tadin D, Huxlin KR. Beyond blindsight: properties of visual relearning in cortically blind fields. *J Neurosci*. 2014;34: 11652–11664.
3. Huxlin KR, Martin T, Kelly K, et al. Perceptual relearning of complex visual motion after V1 damage in humans. *J Neurosci*. 2009;29:3981–3991.
4. Nyquist JB, Lappin JS, Zhang R, Tadin D. Perceptual training yields rapid improvements in visually impaired youth. *Sci Rep*. 2016;6:37431.
5. Huurneman B, Boonstra FN, Goossens J. Perceptual learning in children with infantile nystagmus: effects on reading performance. *Invest Ophthalmol Vis Sci*. 2016;57:4239–4246.
6. Huurneman B, Boonstra FN, Goossens J. Perceptual learning in children with infantile nystagmus: effects on 2D oculo-

motor behavior. *Invest Ophthalmol Vis Sci*. 2016;57:4229–4238.

7. Huurneman B, Boonstra FN, Goossens J. Perceptual learning in children with infantile nystagmus: effects on visual performance. *Invest Ophthalmol Vis Sci*. 2016;57:4216–4228.
8. Fahle M, Henke-Fahle S. Interobserver variance in perceptual performance and learning. *Invest Ophthalmol Vis Sci*. 1996; 37:869–877.
9. Schmitt C, Kromeier M, Bach M, Kommerell G. Interindividual variability of learning in stereoacuity. *Graefes Arch Clin Exp Ophthalmol*. 2002;240:704–709.
10. Astle AT, Li RW, Webb BS, Levi DM, McGraw PV. A Weber-like law for perceptual learning. *Sci Rep*. 2013;3:1158.
11. Li RW, Klein SA, Levi DM. Prolonged perceptual learning of positional acuity in adult amblyopia: perceptual template retuning dynamics. *J Neurosci*. 2008;28:14223–14229.
12. Yehezkel O, Sterkin A, Lev M, Levi DM, Polat U. Gains following perceptual learning are closely linked to the initial visual acuity. *Sci Rep*. 2016;6:25188.
13. Li RW, Provost A, Levi DM. Extended perceptual learning results in substantial recovery of positional acuity and visual acuity in juvenile amblyopia. *Invest Ophthalmol Vis Sci*. 2007;48:5046–5051.
14. Chang LH, Shibata K, Andersen GJ, Sasaki Y, Watanabe T. Age-related declines of stability in visual perceptual learning. *Curr Biol*. 2014;24:2926–2929.
15. Tsirlin I, Colpa L, Goltz HC, Wong AM. Behavioral training as new treatment for adult amblyopia: a meta-analysis and systematic review. *Invest Ophthalmol Vis Sci*. 2015;56: 4061–4075.
16. Andersen GJ, Ni R, Bower JD, Watanabe T. Perceptual learning, aging, and improved visual performance in early stages of visual processing. *J Vis*. 2010;10(13):4.
17. Astle AT, Blighe AJ, Webb BS, McGraw PV. The effect of normal aging and age-related macular degeneration on perceptual learning. *J Vis*. 2015;15(10):16.
18. Amitay S, Halliday L, Taylor J, Sohoglu E, Moore DR. Motivation and intelligence drive auditory perceptual learning. *PLoS One*. 2010;5:e9816.
19. Goldhacker M, Rosengarth K, Plank T, Greenlee MW. The effect of feedback on performance and brain activation during perceptual learning. *Vis Res*. 2014;99:99–110.
20. Shibata K, Yamagishi N, Ishii S, Kawato M. Boosting perceptual learning by fake feedback. *Vis Res*. 2009;49: 2574–2585.
21. Amitay S, Moore DR, Molloy K, Halliday LF. Feedback valence affects auditory perceptual learning independently of feedback probability. *PLoS One*. 2015;10:e0126412.
22. Dosher BA, Jeter P, Liu J, Lu ZL. An integrated reweighting theory of perceptual learning. *Proc Natl Acad Sci U S A*. 2013; 110:13678–13683.
23. Lin F, Heffner KL, Ren P, Tadin D. A role of the parasympathetic nervous system in cognitive training. *Curr Alzheimer Res*. 2017;14:784–789.
24. Roelfsema PR, van Ooyen A, Watanabe T. Perceptual learning rules based on reinforcers and attention. *Trends Cogn Sci*. 2010;14:64–71.
25. Reavis EA, Frank SM, Tse PU. Caudate nucleus reactivity predicts perceptual learning rate for visual feature conjunctions. *Neuroimage*. 2015;110:171–181.
26. Astle AT, Webb BS, McGraw PV. Can perceptual learning be used to treat amblyopia beyond the critical period of visual development? *Ophthalmic Physiol Opt*. 2011;31:564–573.
27. Seitz A, Watanabe T. A unified model for perceptual learning. *Trends Cogn Sci*. 2005;9:329–334.
28. Seitz AR, Dinse HR. A common framework for perceptual learning. *Curr Opin Neurobiol*. 2007;17:148–153.
29. Karni A, Bertini G. Learning perceptual skills: behavioral probes into adult cortical plasticity. *Curr Opin Neurobiol*. 1997;7:530–535.
30. Herzog MH, Fahle M. Effects of biased feedback on learning and deciding in a vernier discrimination task. *Vis Res*. 1999; 39:4232–4243.
31. Keiflin R, Janak PH. Dopamine prediction errors in reward learning and addiction: from theory to neural circuitry. *Neuron*. 2015;88:247–263.
32. Stahl SM. *Essential Psychopharmacology*. 2nd ed. New York: Cambridge University Press; 2003:601.
33. Donovan I, Szpiro S, Carrasco M. Exogenous attention facilitates location transfer of perceptual learning. *J Vis*. 2015;15(10):11.
34. Goldstone RL. Perceptual learning. *Annu Rev Psychol*. 1998; 49:585–612.
35. Watanabe T, Sasaki Y. Perceptual learning: toward a comprehensive theory. *Annu Rev Psychol*. 2015;66:197–221.
36. Haase W, Hohmann A. A new test (C-test) for quantitative examination of crowding with test results in amblyopic and ametropic patients (author's transl). *Klin Monbl Augenheilk*. 1982;180:210–215.
37. Huurneman B, Boonstra FN, Cillessen AH, van Rens G, Cox RF. Crowding in central vision in normally sighted and visually impaired [corrected] children aged 4 to 8 years: the influence of age and test design. *Strabismus*. 2012;20:55–62.
38. Mortimore P, Sammons P, Stoll L, Lewis D, Ecob R. *School Matters: The Junior Years*. Wells, Somerset: Open Books; 1988.
39. Davies JB, Brember I. The reliability and validity of the "Smiley" scale. *Br Educ Res J*. 1994;20:447–454.
40. Kleiner M, Brainard D, Pelli D, Ingling A, Murray R, Broussard C. What's new in Psychtoolbox-3. *Perception*. 2007;36:1.
41. Jeon ST, Hamid J, Maurer D, Lewis TL. Developmental changes during childhood in single-letter acuity and its crowding by surrounding contours. *J Exp Child Psych*. 2010;107:423–437.
42. Norgett Y, Siderov J. Foveal crowding differs in children and adults. *J Vis*. 2014;14(12):23.
43. Gottlob I. Albinism: a model of adaptation of the brain in congenital visual disorders. *Br J Ophthalmol*. 2007;91:411–412.
44. Abadi RV, Bjerre A. Motor and sensory characteristics of infantile nystagmus. *Br J Ophthalmol*. 2002;86:1152–1160.
45. Lee H, Sheth V, Bibi M, et al. Potential of handheld optical coherence tomography to determine cause of infantile nystagmus in children by using foveal morphology. *Ophthalmology*. 2013;120:2714–2724.
46. von dem Hagen EA, Houston GC, Hoffmann MB, Jeffery G, Morland AB. Retinal abnormalities in human albinism translate into a reduction of grey matter in the occipital cortex. *Eur J Neurosci*. 2005;22:2475–2480.
47. Bridge H, von dem Hagen EA, Davies G, et al. Changes in brain morphology in albinism reflect reduced visual acuity. *Cortex*. 2014;56:64–72.
48. Neveu MM, von dem Hagen E, Morland AB, Jeffery G. The fovea regulates symmetrical development of the visual cortex. *J Comp Neurol*. 2008;506:791–800.
49. Jiang J, Zhu W, Shi F, et al. Thick visual cortex in the early blind. *J Neurosci*. 2009;29:2205–2211.
50. Leguire LE, Kashou NH, Fogt N, et al. Neural circuit involved in idiopathic infantile nystagmus syndrome based on FMRI. *J Pediatr Ophthalmol Strabismus*. 2011;48:347–356.
51. Thomas MG, Crosier M, Lindsay S, et al. The clinical and molecular genetic features of idiopathic infantile periodic alternating nystagmus. *Brain*. 2011;134:892–902.

52. Levi DM, Li RW. Perceptual learning as a potential treatment for amblyopia: a mini-review. *Vis Res*. 2009;49:2535–2549.
53. Plude DJ, Enns JT, Brodeur D. The development of selective attention: a life-span overview. *Acta Psychol (Amst)*. 1994;86: 227–272.
54. Petrov AA, Dosher BA, Lu ZL. The dynamics of perceptual learning: an incremental reweighting model. *Psychol Rev*. 2005;112:715–743.
55. Hsu NS, Jaeggi SM. The emergence of cognitive control abilities in childhood. *Curr Top Behav Neurosci*. 2014;16: 149–166.
56. Deveau J, Seitz AR. Applying perceptual learning to achieve practical changes in vision. *Front Psychol*. 2014;5:1166.
57. Fahle M. Perceptual learning: specificity versus generalization. *Curr Opin Neurobiol*. 2005;15:154–160.
58. Jeter PE, Dosher BA, Liu SH, Lu ZL. Specificity of perceptual learning increases with increased training. *Vis Res*. 2010;50: 1928–1940.
59. Ahissar M, Hochstein S. Task difficulty and the specificity of perceptual learning. *Nature*. 1997;387:401–406.
60. Murphy G. Affect and perceptual learning. *Psychol Rev*. 1956; 63:1–15.
61. Bilger M, Manning WG. Measuring overfitting in nonlinear models: a new method and an application to health expenditures. *Health Econ*. 2015;24:75–85.