nature neuroscience

Perceptual learning in autism: over-specificity and possible remedies

Hila Harris[1], David Israeli[2], Nancy Minshew[3], Yoram Bonneh[1,4], David J Heeger[5], Marlene Behrmann[6,7] & Dov Sagi[1,7]

Inflexible behavior is a core characteristic of autism spectrum disorder (ASD), but its underlying cause is unknown. Using a perceptual learning protocol, we observed initially efficient learning in ASD that was followed by anomalously poor learning when the location of the target was changed (over-specificity). Reducing stimulus repetition eliminated over-specificity. Our results indicate that inflexible behavior may be evident ubiquitously in ASD, even in sensory learning, but can be circumvented by specifically designed stimulation protocols.

Cognitive inflexibility[1] and atypical learning[2] are key characteristics of the ASD phenotype; however, little is known about the mechanisms that underlie this behavior and about procedures that might circumvent the abnormally restricted learning. It is also unclear whether restricted behaviors are unique to the social-cognitive aspect of ASD or whether they also affect sensory functions. We characterized learning in ASD using an established protocol in which the perceptual learning parameters of typically developing (TD) observers are well documented[3]. Mainly, it has been shown that over-repetitive stimulation slows down perceptual learning and increases specificity. When the detrimental effect of stimulus repetition is attenuated, however, generalization is achieved[4].

High-functioning ASD adults and matched control observers were trained with a texture discrimination task[5] consisting of a target display followed by a patterned mask (**Fig. 1**). Observers were required to judge the orientation of a peripheral target consisting of three diagonal bars surrounded by horizontal lines. Performance was measured as a function of the time interval between the onset of the target and the mask (stimulus onset asynchrony, SOA), with threshold defined as the minimal time (SOA) to reach a predefined criterion level of performance (Online Methods). Observers completed four daily sessions with the target situated in location 1. Generalization of learning across target locations was assessed by four subsequent sessions (days 5–8) with the target appearing at location 2.

A group of ASD observers (ASD standard) and their control group (control standard) were trained with this standard training protocol. Although substantial ASD learning was achieved during location 1 training (**Fig. 2a**, see Online Methods for statistical analysis), performance at the new location (day 5) was poorer than the performance level at the previous location (day 4), with the threshold equivalent to that measured initially on day 1. This reflects specificity of learning. Furthermore, additional training at location 2 resulted in slower learning relative to that observed for location 1 (**Fig. 2b**), indicating over-specificity of learning. This over-specificity, observed here for the first time, to the best of our knowledge, contrasts with the established perceptual learning profile[3] of TD observers who evince faster learning at location 2 than at location 1. Transfer cost (TC), quantified as the average difference between the transfer thresholds and their corresponding training thresholds (**Fig. 2c** and Online Methods), captures the difference between the learning curve in the new location relative to the learning curve at the original location. Positive values, indicating over-specificity, reflect hampered ability to learn a new condition following prior learning. This was clearly the case for the ASD standard group (**Fig. 2c**), whose TC was substantially higher than that of the matched control group. This difference reflects the unique over-specificity in the ASD group (**Fig. 2** and **Supplementary Fig. 1**). The control standard group showed negative TC, as expected from the literature[3], indicating faster learning in location 2 relative to location 1. However, learning for this group was non-specific, as indicated by the threshold on day 5. Given that this group demonstrated weak learning at location 1 (28 ± 7 ms, $P = 0.05$), it is difficult to determine whether performance on day 5 is truly reflective of generalization or of partial transfer[3]. This behavior might be explained by the easier conditions[6] (such as long target duration) adopted here (Online Methods) to ensure that ASD observers could perform the task.

A second group of ASD observers and their matched control group completed the same training protocol, but with 'dummy' trials consisting of the textured background and no target, interleaved

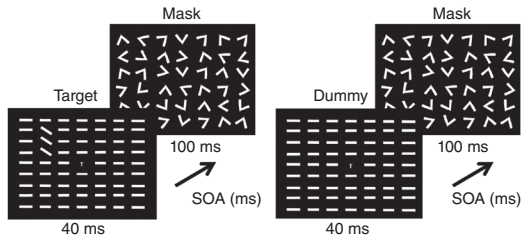

Figure 1 The texture discrimination task. Schematic illustration of standard trials (left). The target frame consists of three diagonal bars differing in orientation from a background of horizontal identical bars. After a blank inter-stimulus interval (SOA), a mask frame appears. Observers indicate, using the computer mouse keys, whether a small rotated letter at the center of the display is "T" or "L" (fixation target) and whether the three-bar arrangement is vertical or horizontal (Online Methods). Schematic illustration of dummy trials is shown on the right.

[1]Department of Neurobiology/Brain Research, Weizmann Institute of Science, Rehovot, Israel. [2]Jerusalem Center for Mental Health, Hebrew University, Jerusalem, Israel. [3]Center for Excellence in Autism Research, University of Pittsburgh, Pittsburgh, Pennsylvania, USA. [4]Department of Human Biology, University of Haifa, Haifa, Israel. [5]Department of Psychology and Center for Neural Science, New York University, New York, New York, USA. [6]Department of Psychology, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA. [7]These authors contributed equally to this work. Correspondence should be addressed to D.S. (dov.sagi@weizmann.ac.il).

Received 29 July; accepted 8 September; published online 5 October 2015; doi:10.1038/nn.4129

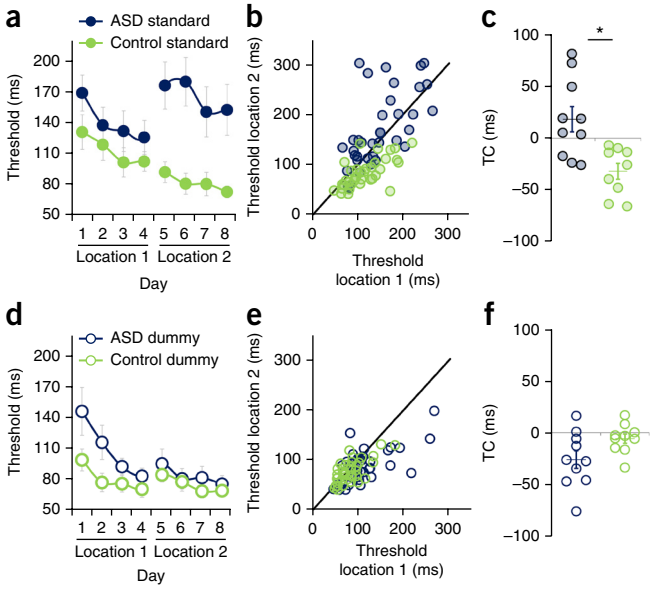

Figure 2 Over-specificity in ASD visual learning and its elimination following the dummy training. (**a**) Learning curves for standard training. Blue indicates the ASD group ($n = 10$; day 1, 170 ± 17 ms; day 4, 127 ± 16 ms; Student's two-tailed t test, $P = 0.01$; day 5, 177 ± 23 ms; $P = 0.016$ compared with day 4) and green indicates the matched control group ($n = 9$; $P = 0.05$ between days 1–4, $P = 0.12$ between days 4–5). (**b**) Thresholds at location 2 (days 5–8) versus the corresponding thresholds (days 1– 4) at location 1. Each data point represents the thresholds for one individual observer on corresponding days (for example, day 5 versus day 1 or day 8 versus day 4). The black line is the identity line. (**c**) Individual TC of learning for ASD and matched controls (corresponding to the mean difference between the ordinate and abscissa values of the data points in **b**; ASD standard, 18 ± 12 ms d^{-1}; control standard, −32 ± 8 ms d^{-1}; one-way ANOVA, $^*P = 0.003$). Error bars represent s.e.m. (**d**–**f**) Learning curves for dummy training. Data are presented as in **a**–**c**, but for the groups of observers tested with the dummy trials interleaved (ASD, $n = 10$; control, $n = 10$; $P < 0.01$ between days 1–4, $P = 0.54$ between days 4–5, $P = 0.05$ for the TC comparison).

randomly with the standard trials. We hypothesized that the insertion of these dummy stimuli would result in generalization by reducing the detrimental effect of stimulus repetition[4,7] (Online Methods).

Both dummy groups (ASD and control) showed generalization of learning following training (**Fig. 2d**–**f**). The ASD dummy group achieved better performance at the end of training relative to the ASD standard group (84 ± 8 ms versus 127 ± 16 ms; $P = 0.03$), indicating that reducing the effect of stimulus repetition facilitated learning in ASD observers. This group subsequently showed (days 5–8) full generalization of learning in addition to negative TC (TC = −26 ± 8 ms). Thus, unlike the ASD standard group, the ASD dummy group did not show over-specificity and exhibited transfer of learning. Moreover, their learning profile was comparable to that of the control dummy group (**Fig. 2d**–**f**).

The efficient learning that occurred when the target was in location 1 is consistent with findings of intact implicit[8] and statistical[9] learning in ASD. What has not been seen previously is the inflexibility to a contextual change (location). This inflexibility cannot be attributed to an overall difficulty of the ASD observers in performing the task[6]; in location 1, the same performance difference was shown between the ASD and control across standard (27 ± 4 ms) and dummy (28 ± 8 ms) training conditions. Notably, this atypical over-specificity of learning in ASD could be avoided by attenuating the effects of stimulus repetition. Thus, the dummy training method, previously shown to reduce specificity in TD observers, serves as a useful procedure in counteracting over-specificity in ASD.

In accordance with previous findings[10], reaction times (RTs) of our TD observers showed generalization of RT gains across locations (**Fig. 3**). This suggests that RT improvement is mainly explained by general factors, such as motor response, information transmission time or learning the temporal properties of the stimulus[10,11]. However, ASD observers trained in the standard condition do not improve at the new location beyond the level achieved at day 4. This was evident by the absence of RT gains from days 4–8 (**Fig. 3a**,**b**). ASD observers trained with dummy stimuli did improve in location 2 (**Fig. 3c**,**d**), similar to the matched control group. The ASD dummy group showed a training-independent slowdown (274 ms, s.d. = 38 ms) that might be attributable to the randomly mixed dummy trials. Overall, the reaction times analysis supports over-specificity in the ASD standard group, but not in the ASD dummy group.

These findings have both theoretical and practical implications for our understanding of ASD. In the field of perceptual learning, specificity and the experimental parameters affecting it have been extensively documented with typically developing observers[3]. According to learning theories, extended learning with a fixed target is expected to result in narrow learning, restricted to the exact trained target (as a result of over-fitting; see refs. 3,12). We found a lack of generalization when the same trained target was moved to a new location. This suggests that location specificity of learning depends on spatial heterogeneity in early visual representations[3,4]. With repetitions,

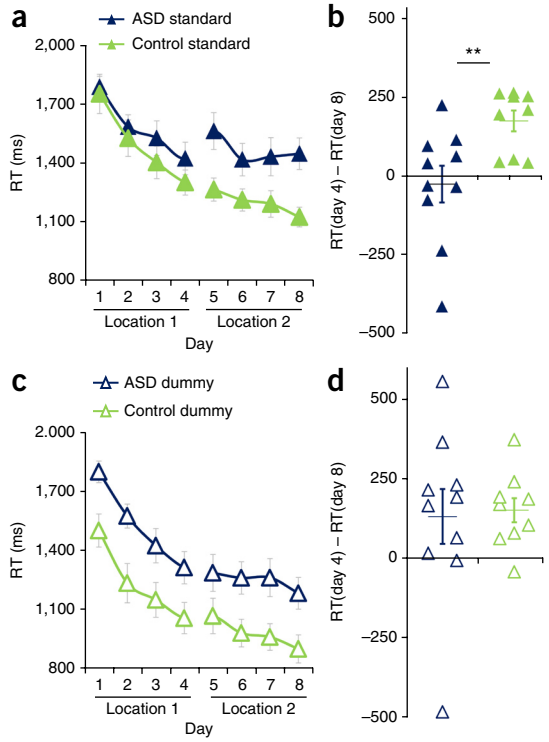

Figure 3 RT analysis. (**a**,**c**) RT during training for the standard (ASD, $n = 10$; control, $n = 9$) and dummy (ASD, $n = 10$; control, $n = 9$) groups. (**b**,**d**) RT gains from days 4–8 for the standard groups (ASD, −24 ± 58 ms; control, 175 ± 33; two-sample t test, $^{**}P = 0.01$) and dummy groups (ASD, 130 ± 86 ms; control, 150 ± 38 ms; $P = 0.08$). Error bars represent s.e.m.

effects such as sensory adaptation[13] (reduced sensitivity following repetitions) may cause distinct subpopulations of neurons (for example, those that respond to location 1 versus those that respond to location 2) to encode stimuli differently[5]. Over-specificity in ASD may be a consequence of such heterogeneity in the visual cortex, possibly combined with a failure of higher levels of processing to handle the perturbed input. The dummy trials were effective in both ASD and controls, revealing that reduced sensory adaptation promotes greater spatial invariance in the neural representations of the stimuli[4,14].

Our results suggest that repetition, a technique widely used in intervention and education for ASD, may lead to inflexibility. The adverse consequences of repetition may apply to an even greater degree as the complexity of learning and behavior increases, such as in the domain of social behavior. Counterintuitively, reducing stimulus repetition may enhance learning and foster generalization in ASD.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank R. Egan, A. Gupta and J. Geskin for their assistance in data collection, and D. Plaut and R. Malach for fruitful comments and discussions. This work was supported by grants from the US-Israel Binational Science Foundation and the Simons Foundation Autism Research Initiative. H.H. thanks the Azrieli Foundation for the award of an Azrieli Fellowship.

AUTHOR CONTRIBUTIONS

H.H., M.B. and D.S. designed the experiments. D.I., N.M. and Y.B. handled the clinical aspects. H.H. and M.B. collected the data. H.H., M.B. and D.S. analyzed the data. H.H., D.J.H., M.B. and D.S. wrote the paper.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Sumiyoshi, C. *et al. Neurosci. Res.* **69**, 252–257 (2011).
2. Dawson, M., Mottron, L. & Gernsbacher, M.A. in *Learning and Memory: a Comprehensive Reference* (eds. Byrne, J. & Roediger, H.L.I.) 759–772 (Elsevier, Oxford, 2008).
3. Sagi, D. *Vision Res.* **51**, 1552–1566 (2011).
4. Harris, H., Gliksberg, M. & Sagi, D. *Curr. Biol.* **22**, 1813–1817 (2012).
5. Karni, A. & Sagi, D. *Proc. Natl. Acad. Sci. USA* **88**, 4966–4970 (1991).
6. Ahissar, M. & Hochstein, S. *Nature* **387**, 401–406 (1997).
7. Greenlee, M.W. & Magnussen, S. *Vision Res.* **28**, 1303–1310 (1988).
8. Brown, J., Aczel, B., Jiménez, L., Kaufman, S.B. & Grant, K.P. *Q. J. Exp. Psychol. (Hove)* **63**, 1789–1812 (2010).
9. Roser, M.E., Aslin, R.N., McKenzie, R., Zahra, D. & Fiser, J. *Neuropsychology* **29**, 163–172 (2015).
10. Petrov, A.A., Van Horn, N.M. & Ratcliff, R. *Psychon. Bull. Rev.* **18**, 490–497 (2011).
11. Liu, C.C. & Watanabe, T. *Vision Res.* **61**, 107–114 (2012).
12. Qian, N. & Lipkin, R.M.A. *Front. Hum. Neurosci.* **5**, 77 (2011).
13. Webster, M.A. *J. Vis.* **11**, 3 (2011).
14. Harris, H. & Sagi, D. *Vision Res.* **109**, 77–86 (2015).

ONLINE METHODS

Subjects. 23 high-functioning observers with ASD (1 female, mean age of 26, range of 19–41 years) and 19 age- and gender-matched controls (mean age of 28, range of 24–35 years) participated in this study. All observers had normal or corrected-to-normal vision and provided written informed consent. Carnegie Mellon University Institutional and the Weizmann Institute of Science Review Board approved the protocol. The diagnosis of autism was established using the Autism Diagnostic Observation Schedule-G[15], and confirmed by expert clinical diagnosis (N. Minshew and D. Israeli) and, in Pittsburgh, by the additional use of the Autism Diagnostic Interview-Revised[16]. The ASD individuals were medically healthy, had no identifiable genetic, metabolic or infectious etiology for their disorder, and had no history of traumatic brain injury, seizures, attention deficit disorder or depression. IQ was assessed in the ASD individuals using the Weschler Abbreviated Scales of Intelligence (WASI) and all ASD participants had Full Scale IQ scores above 85. About 50% of all ASD individuals score 85 and above on IQ measures[15]; thus, examining learning in this high-functioning ASD sample is likely to be representative of at least half of the ASD population. The observers were randomly assigned to the standard or dummy group. Subsets of participants completed different training methods, but each group completed 8 d of testing and each individual participant yielded large amounts of data for fine-grained analysis. Two data points of one ASD observer were excluded due to the onset of medical treatment. One observer from the control dummy group completed only 6 d of training due to schedule constrains, thus his data was excluded from the RT analysis.

The data of the non-matched control group shown in **Supplementary Figure 1** is reproduced from ref. 4.

Stimuli and procedure. Observers were tested individually in a quiet room. They were seated comfortably in front of a computer screen. A display appeared on the screen and the participant reported whether a "T" or an "L" appeared at the center of the display (**Fig. 1**). Next, the participant reported whether the texture target (array of three peripheral bars embedded in the background) was oriented horizontally or vertically. Responses were provided by pressing one of two preassigned mouse keys for each of the two reports (first response: left click "L", right click "T"; second response: left click "horizontal", right click "vertical"). For example, for a stimulus presenting a "T" and "Horizontal" target the correct response would be one right click followed by one left click. Auditory feedback was provided to indicate an incorrect response for the fixation (T/L) task only (no feedback was provided for the peripheral horizontal/vertical texture task). Performance on the texture discrimination task was measured as a function of the time-interval between target and mask onsets (SOA, ranging from target duration to 800 ms). The SOA was randomized across trials. Based on the results of a pilot study, we modified several parameters relative to our previously reported protocol[4] to ensure that the ASD individuals would be able to complete the task well: the target presentation time was 40 ms (instead of only 10 ms), fixation target (T/L) was slightly enlarged, and two easier SOA's were added to the SOA range in the training phase. Each trial was self-initiated by the observer, resulting in ~2-s intertrial interval. Discrimination thresholds were estimated by fitting a Weibull function to the psychometric data (performance accuracy at the peripheral task versus SOA)[4]. In the dummy condition, background-only stimuli (no texture target) were randomly interleaved with the test trials (**Fig. 1**), but observers responded to both the fixation target and the texture target on all trials, guessing the second response if not detecting the presence of a texture target. The dummy trials were hypothesized to minimize adaptation to the target since the background horizontal bars presented on those trials were oriented 45° relative to the targets' local orientation on the standard trials[4,7]. Each observer participated in a number of pre-training trials on day 1. The number of trials during this pre-training phase was adjusted for each observer. The criterion for this pre-training phase was 100% correct for a short session of 10 trials at a high and constant SOA (800 ms), and this pre-defined performance criterion was selected to ensure that all participants started the experiment at the same level of performance. On day 5, observers were verbally informed about the change in location but no pre-training was performed at the untrained location. Each daily session consisted of 288 target trials with an additional 288 trials in the dummy condition.

The TC was quantified as the average difference between the transfer thresholds and their corresponding training thresholds (specifically, Th_{day5}–Th_{day1}, Th_{day6}–Th_{day2}, Th_{day7}–Th_{day3}, Th_{day8}–Th_{day4}).

The observers were requested to respond as accurately as they can. The RTs of the peripheral texture task (horizontal/vertical) were analyzed. The daily RT median was calculated for each observer. Only RTs in the range of 400–4,000 ms were included in the RT analysis (98.7% of total trials). The across-observers means of these daily medians are plotted in **Figure 3**.

Statistical analysis. A three-way repeated measures ANOVA was used to reveal a significant effect of day ($F_{(3,99)} = 22$, $P < 0.001$), and a significant day × group (ASD/control) × condition (standard/dummy) interaction ($F_{(3,99)} = 5.8$, $P = 0.003$). Non-significant interactions were found for day × group ($F_{(3,99)} = 2.7$, $P = 0.06$) and for day × condition ($F_{(3,99)} = 0.7$, $P = 0.5$).

Figure 2, upper panel, standard training. A two-way repeated-measures ANOVA (Greenhouse-Geisser) revealed a significant effect of day ($F_{(3,48)} = 7.7$, $P = 0.001$) and a significant day × group interaction ($F_{(3,48)} = 5.2$, $P = 0.008$). Subsequently, the effect of day was examined for each group independently. One way-ANOVA with repeated-measures found a significant effect of day in each group (ASD, $F_{(3,24)} = 4.8$, $P = 0.02$; control, $F_{(3,24)} = 9.9$, $P = 0.003$). A subsequent pairwise t test comparison (two tailed, Bonferroni corrected) revealed that the ASD standard group improved significantly during training (from day 1 to day 4, $t(9) = 3$, $P = 0.01$) unlike the control standard group in which minimal improvement was shown ($t(8) = 2.2$, $P = 0.05$). Of interest, whereas the ASD standard' group showed significant specificity of learning (poorer thresholds day 5 relative to day 4, $t(9) = 2.9$, $P = 0.016$), this was not true of the control standard group ($t(8) = -1.7$, $P = 0.12$). The TC was compared using one-way ANOVA showing a significant effect of group ($F_{(1,17)} = 11.5$, $P = 0.003$).

Figure 2, lower panel, dummy training. A two-way repeated-measures ANOVA (Greenhouse-Geisser corrected) revealed a significant effect of day ($F_{(3,51)} = 16.4$, $P < 0.001$) and a non-significant day × group interaction ($F_{(3,51)} = 2.8$, $P = 0.09$). Subsequently, the effect of day was examined, revealing significant learning (from day 1 to day 4, $t(19) = 4.6$, $P < 0.01$, Bonferroni corrected), and generalization (day 4 relative to day 5, $t(19) = 2$, $P = 0.54$). The TC was compared using one-way ANOVA showing a non-significant group effect ($F(1,18) = 4.4$, $P = 0.05$). The distribution of the data was assumed to be normal, but this was not formally tested. Sphericity violations were corrected using Greenhouse-Geisser procedure. No randomization or blinding was employed during data analysis.

Figure 3, standard training. In order to compare the RT gains from day 4 to day 8 across ASD and control groups, a two-sample t test (two-tailed) was performed, revealing a significant difference ($t(17) = -2.9$, $P = 0.01$).

Figure 3, dummy training. In order to compare the RT gains from day 4 to day 8 across ASD and control groups, a two- sample t test (two-tailed) was performed, revealing a non-significant difference ($t(17) = -0.2$, $P = 0.8$).

No statistical methods were used to pre determine sample size but our sample sizes are similar to those generally employed in the field.

A **Supplementary Methods Checklist** is available.

15. Lord, C. *et al. J. Autism Dev. Disord.* **30**, 205–223 (2000).
16. Lord, C. *et al. J. Autism Dev. Disord.* **24**, 659–685 (1994).