



Adapted to explore: Reinforcement learning in Autistic Spectrum Conditions

Eldad Yechiam^{a,*}, Olga Arshavsky^a, Simone G. Shamay-Tsoory^b, Shoshana Yaniv^b, Judith Aharon^b

^aTechnion – Israel Institute of Technology, Haifa, Israel

^bHaifa University and Rambam Medical Center, Haifa, Israel

ARTICLE INFO

Article history:

Accepted 15 October 2009

Available online 12 November 2009

Keywords:

Autism
Reinforcement
Learning
Modeling
Decision making

ABSTRACT

Recent studies have recorded a tendency of individuals with Autism Spectrum Conditions (ASC) to continually change their choices in repeated choice tasks. In the current study we examine if this finding implies that ASC individuals have a cognitive style that facilitates exploration and discovery. Six decision tasks were administered to adolescents with ASC and matched controls. Significant differences in shifting between choice options appeared in the Iowa Gambling task (Bechara, Damasio, Damasio, & Anderson, 1994). A formal cognitive modeling analysis demonstrated that for about half of the ASC participants the adaptation process did not conform to the standard reinforcement learning model. These individuals were only coarsely affected by choice-outcomes, and were more influenced by the exploratory value of choices, being attracted to previously un-explored alternatives. An examination of the five simpler decision tasks where the advantageous option was easier to determine showed no evidence of this pattern, suggesting that the shifting choice pattern is not an uncontrollable tendency independent of task outcomes. These findings suggest that ASC individuals have a unique adaptive learning style, which may be beneficial in some learning environment but maladaptive in others, particularly in social contexts.

© 2009 Elsevier Inc. All rights reserved.

1. Introduction

In 2005 Nobel laureate Vernon Smith, considered the father of experimental economics, openly discussed his diagnosed Asperger's syndrome, and suggested that it helps improve his ability to make scientific discoveries. Asperger's syndrome is a mild form of autism,¹ which along with high functioning autism, and pervasive developmental disorder are known as Autistic Spectrum Conditions (ASC) (Baron-Cohen, 2009; Golan, Baron-Cohen, Hill, & Rutherford, 2007).

Researchers have suggested that many other well-known inventors and discoverers, such as Issac Newton and the mathematician Paul Erdos, had ASC (Fitzgerald, 2004). We examined whether adolescents with ASC indeed have an adaptive learning style geared towards exploration and discovery.

Recent studies revealed a unique decision making pattern in ASC, characterized by constant shifting between choice alternatives in repeated choice tasks (Johnson, Yechiam, Murphy, Queller, & Stout, 2006; Minassian, Paulus, Lincoln, & Perry, 2006). For

example Johnson et al. (2006) examined the behavior of adolescents with ASC on the Iowa Gambling task (Bechara et al., 1994), a repeated choice task involving four decks of cards, in which two of the decks are advantageous and two are disadvantageous (see Section 2 and Table 1 for a complete description of the task). The individual learning curves of adolescents with ASC revealed that they tended to constantly shift from one choice alternative to the other, and this pattern did not diminish with task experience. The difference in this behavior was evaluated by calculating the average and maximal size of consecutive selections (or runs) from the same alternative. In the control group, the average run of consecutive choices was about three trials on average (implying that a person selected a deck three times in a row). In contrast, in the ASC group it was only 1.3 trials (i.e., most selections were followed by a switch). Similar differences emerged for the largest run of consecutive choices. Johnson et al. (2006) also found that within the ASC group short runs on the IGT were associated with more severe autistic syndromes on the Autism Diagnostic Interview (ADI-Revised, Lord, Rutter, & Le Couteur, 1994). Fig. 1 shows an example of this exceptional adaptive learning style in three of our study participants diagnosed with ASC (left) and three matched healthy controls (right). As one can see, the ASC participants' adaptation pattern involves constant flipping between choice alternatives. In Johnson et al.'s (2006) study this pattern (of an average run of less than 1.5) characterized 80% of the adolescents diagnosed with ASC.

Using formal reinforcement learning models we contrasted three possible explanations of this pattern. The first explanation

* Corresponding author. Address: Behavioral Science Area, Faculty of Industrial Engineering and Management, Technion, Haifa 32000, Israel. Fax: +972 4 829 5688. E-mail address: yeldad@tx.technion.ac.il (E. Yechiam).

¹ Asperger's syndrome is distinct in that language is not impaired, and IQ is often in the superior range (Fitzgerald & Corvin, 2001). In addition, as opposed to other Autism Spectrum Conditions, diagnosis is often not made until school age, implying that most children with Asperger's syndrome have not experienced an early intensive intervention (which could potentially influence learning style) (Baird et al., 2001).

Table 1
The payoffs in the experimental tasks: the Iowa Gambling task (IGT) and the Simplified Binary tasks (SBT).

Task	Card	Outcome	Description
IGT	A	Win 1.0 every card and .5 to loss 2.5	Disadvantageous
	B	1.0 every card and .1 to loss 12.5	Disadvantageous
	C	0.5 every card and .5 to loss 0.5	Advantageous
	D	0.5 every card and .1 to loss 2.5	Advantageous
SBT1	H	5 every card	Advantageous
	L	0 every card	Disadvantageous
SBT2	H	.5 to gain 4 and .5 to gain 5	Advantageous
	L	.5 to gain 2 and .5 to gain 3	Disadvantageous
SBT3	H	.5 to gain 2 and .5 to gain 6	Advantageous
	L	.5 to gain 2 and .5 to gain 3	Disadvantageous
SBT4	H	.33 to gain 5, .33 to gain 6, .33 to lose 1	Advantageous
	L	.33 to gain 4, .33 to gain 5, .33 to lose 1	Disadvantageous
SBT5	H	.33 to gain 1, .33 to lose 4, .33 to lose 5	Advantageous
	L	.33 to gain 1, .33 to lose 5, .33 to lose 6	Disadvantageous

is that the choices of ASC individuals are simply more random, due to errors in selection. This is consistent with the idea that complex tasks that require generalization skills, such as the IGT, pose a challenge to people with ASC because of their difficulty in recognizing relationships between features of the task and forming general

knowledge about categories of items and types of situations (see Klinger & Dawson, 2001; Klinger, Klinger, & Pohlig, 2006).

A second explanation is that the adaptation process of individuals with ASC involves more trial and error exploration. This argument is consistent with findings showing that adolescents with ASC rate high on such traits as perseverance and drive for perfection (e.g., Ashburner, Ziviani, & Rodger, 2009; Gillberg, 2002; Kobayashi & Murata, 1998). Under a third, related, explanation the difference in exploration capacity is qualitative rather than quantitative; that is, while individuals with ASC explore more than healthy individuals do, their exploration does not conform to the standard reinforcement learning paradigm. Rather, exploration may be the primary directive of their adaptation process. For examining this last possibility, a conventional model of reinforcement learning was contested with two novel models resting on the assumption that exploration, and not the pursuit of outcomes, is the main directive governing the learning process.

In order to examine the scope of the difference between adolescents with ASC and healthy controls, we employed a battery of decision tasks ranging in their difficulty (see Table 1). The battery included the more complex Iowa Gambling task (Bechara et al., 1994) in which the shifting choice pattern was originally found (Johnson et al., 2006). It also included five simpler variants of this decision tasks. This enabled us to assess if the tendency to continuously shift choices in ASC emerges only in relatively difficult deci-

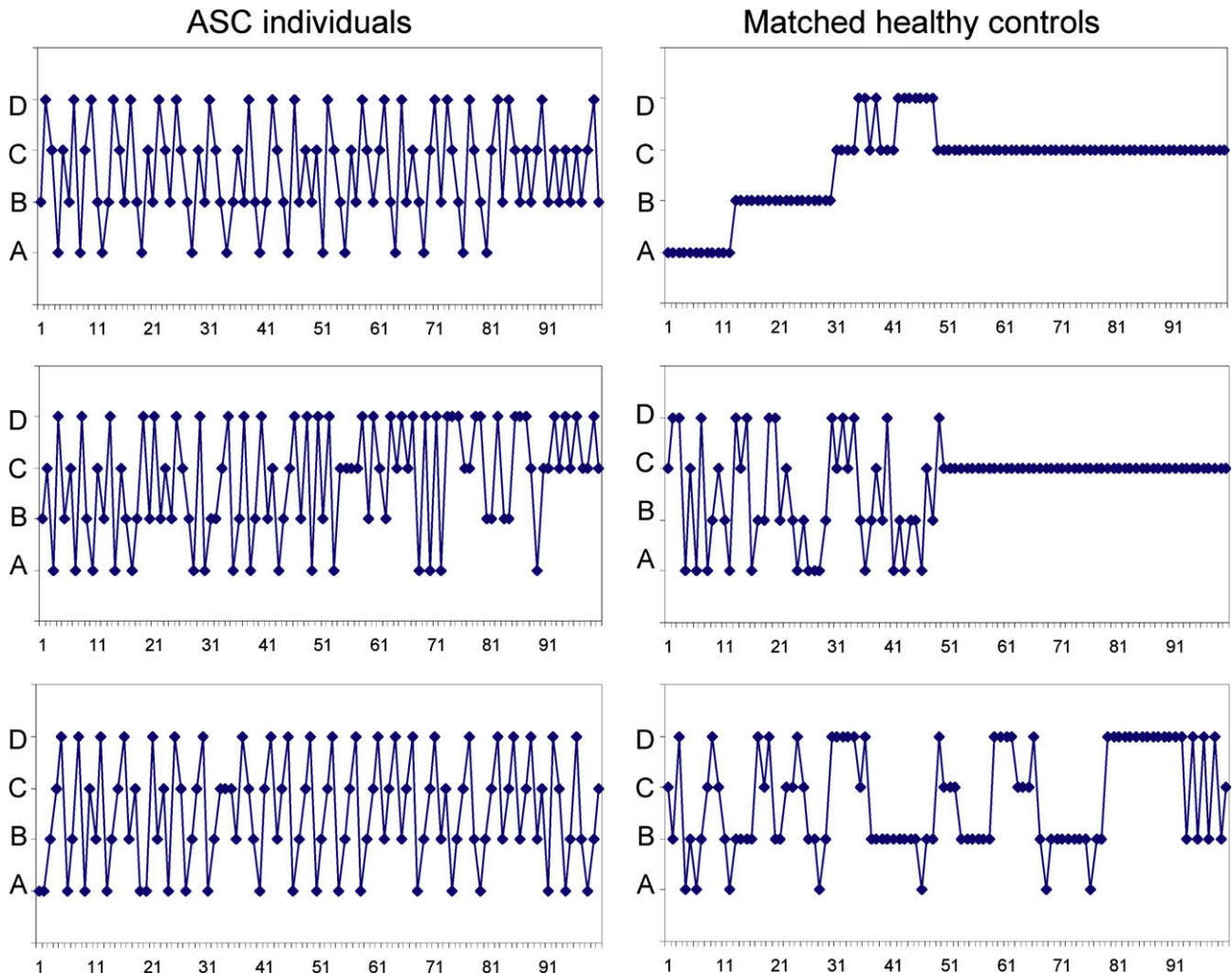


Fig. 1. The Iowa Gambling task performance of three participants with Autistic Spectrum Conditions (ASC) in our study (left) and their matched healthy controls (right). The results from the four decks (A, B, C, and D) in each of 100 trials are shown.

sion tasks, or whether it is an uncontrollable tendency appearing even in simpler tasks. All tasks were then analyzed with the three classes of formal reinforcement learning model noted above. These models are explained more thoroughly in Section 3 and their mathematical details are available in the Appendix section.

2. Method

2.1. Participants

The participants included 15 high functioning children and adolescents with formal diagnoses of ASC, who arrived to the community medical center for diagnosis, either individually or through ads placed in a patient support group center. The majority of the participants in this group (14 out of the 15) were male. The mean group age was 15.6 (SD = 2.8). Only participants diagnosed by at least one psychiatrist as specifically meeting the ICD-10 criteria for two autism spectrum disorders were included: Asperger's syndrome ($n = 12$) and pervasive developmental disorder ($n = 2$). One high functioning participant with autism who had ample language capabilities was also included. In addition, all were screened for the current study using the Autism Spectrum Quotient (AQ) (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001) and only those scoring above the cutoff of 32 (as determined by Baron-Cohen et al., 2001) were included (except for two participants with a clear diagnosis of Asperger's syndrome as suggested by two psychiatrists, whose AQ scores were 13 and 30).

Participants in the control group were recruited by posting ads at local community billboards. Thirty healthy control (HC) participants (28 males and 2 females) were matched by age and gender to the ASC groups. The matching by age ensured that for each participant in the ASC group there were two participants in the same age (± 1 year) in the HC sample. The data from two participants in the HC sample was lost due to a technical error, so this group eventually included only 28 participants (26 males and 2 females). The mean group age in the HC group was 15.6 (SD = 3.6). The final age range in the two groups was 9–21.

Three of the participants in the ASC group were also diagnosed with OCD, and three with ADHD (total of five participants with comorbid disorders, as one individual had both OCD and ADHD in addition to AS). We included these participants in our main analysis as ASC are very often accompanied by secondary disorders. However, to rule out possible effects of these co-morbidities we also separately examined the ASC group without these participants.

2.2. Measures

2.2.1. The Iowa Gambling task (IGT)

A computerized version of the task described in Bechara et al. (1994) was used. The participant sees four decks of cards on a computer screen labeled A, B, C, and D. Using the mouse, the participant can select a card from any of the four decks. After selecting a card, the participant receives money (the amount is displayed on the screen). The payoffs associated with the four decks are presented in Table 1. A horizontal bar at the bottom of the display shows the cumulative payoffs, and is updated with each trial. A second bar shows the amount given to the participant at the beginning of the task, as a "loan".

The initial loan in our study was of NIS 2000, and this amount was added to the cumulative gains bar. The minimum inter-trial interval was set to 0.5 s, and the task included 100 trials. Participants were given verbal instructions identical to those provided in Johnson et al. (2006).

2.2.2. Simplified Binary tasks

In these simpler variants of the IGT the participant sees two virtual unmarked buttons on the computer screen, and is asked to press one of them in each trial using the mouse. After pressing the button, the participant receives money (the amount appears on the selected button and on a bar below the buttons).

The battery included five tasks. In each task one choice alternative, associated with a random button, was advantageous, and the other choice alternative was disadvantageous. The exact payoffs are described in Table 1. The tasks were designed so that it would be increasingly difficult to find the advantageous choice alternative. In Tasks 1 and 2 there is no overlap between the payoffs of the two alternatives; one alternative is always better than the other. In Task 3 there is some overlap between alternatives, and Tasks 4 and 5 include larger overlaps. The degree of overlap in these tasks is still much smaller than in the IGT: in Tasks 4 and 5 the advantageous alternatives are better 55% of the time while in the IGT the advantageous alternatives produce better payoffs only 5% of the time.

The order of appearance of the five tasks was according to their difficulty (1–5). The minimum inter-trial interval was set to 0.5 s. Each task included 30 trials (a total of 150 trials). The complete instructions appear in Yechiam and Ert (2007).

2.2.3. Additional tests

Participants completed verbal and non-verbal intellectual aptitude tests. The verbal aptitude test was the similarities subscale from the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1981). The non-verbal test was the Raven Standard Progressive Matrices Test, (Raven, 1965). In both tests the version administered was appropriate to the age of the participant.

2.3. Procedure

Participants in the ASC group completed the IGT following a formal diagnosis session at the Rambam medical center. After about a month into the study we began to administer the TB tasks, and thus only nine ASC participants completed them. The control participants and the ASC participants answering the ad were given a fee of NIS 50 (about \$12.5) an hour for their participation. Participants were tested in a single session lasting approximately 2 h, with a break of 15 min.

3. Results

Fig. 2 presents the proportion of selections from the advantageous decks of the IGT in the two groups. As can be seen, although both groups learned to make more advantageous choices with experience, learning was slower in the ASC group. To examine the statistical significance of this pattern we conducted a mixed analysis of variance (ANOVA) with group as a between-subject factor, and task block (four blocks of 25 trials) as a repeated measure. The results revealed a significant group by task-block interaction ($F(3, 108) = 2.81, p < .05, MSE = 0.02$).²

We next examined the average and largest runs of consecutive choices from the same choice alternative. One participant in each of the HC and ASC group were outliers with average run sizes of over three standard deviations above the mean, and these were removed from the analysis. A comparison of the largest and average

² The participants with Asperger's syndrome and those diagnosed with PDD and high functioning autism had a similar proportion of advantageous selections. The Asperger's syndrome group chose the advantageous decks 49.3% of the time, on average, and the other ASC participants chose it 48.4% of the time. The learning curves of the two sub-groups were also similar. For conciseness, this comparison is not detailed.

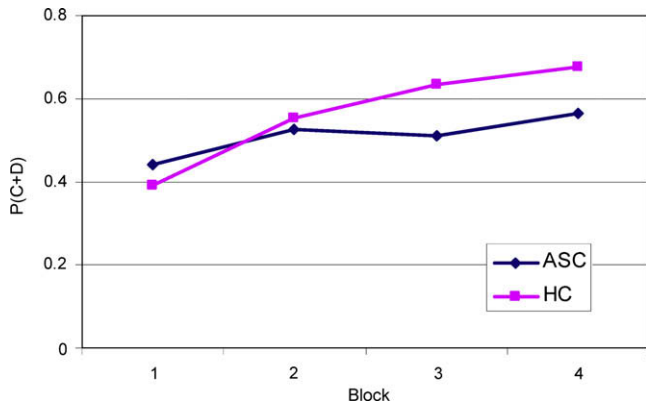


Fig. 2. Mean proportions of selections from the advantageous decks (P(C + D)) of the Iowa Gambling task: a comparison of the ASC and HC groups (four blocks of 25 trials).

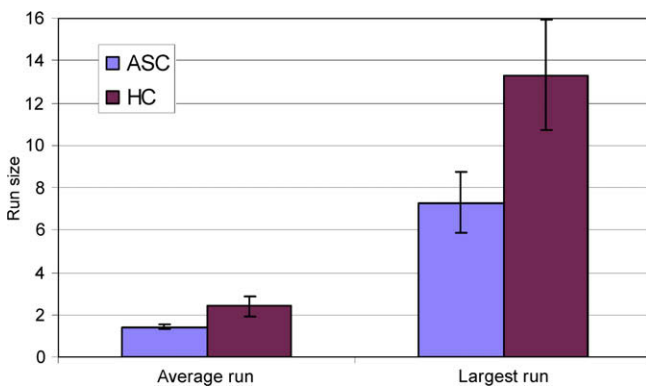


Fig. 3. Largest and average runs of consecutive choices from the same alternative (means and standard errors): a comparison of the ASC and HC groups.

runs for the two groups appears in Fig. 3. Consistent with the pattern observed in Johnson et al. (2006), the average and largest run sizes were about twice as short for the ASC group compared to the control group. In particular, the average run was 1.36 in the ASC group and 2.4 in the HC group, a significant difference ($t(28.5) = 2.09$, $p < .05$; Cohen's $d = 0.58$). The differences in the largest run was also significant ($t(36.7) = 1.97$, $p = .05$ one-tailed; Cohen's $d = 0.37$). An analysis of the individuals in the ASC group who were not diagnosed with Asperger's syndrome indicated they also conformed to the shifting choice pattern (with an average run of 1.15). An analysis of the subset of individuals in the ASC group without co-morbid disorders revealed the same pattern.

3.1. Cognitive modeling analysis

In order to clarify the reason for the short runs exhibited by the ASC group on the IGT, we conducted a cognitive modeling analysis for this task (see Busemeyer & Stout, 2002; Yechiam, Busemeyer, Stout, & Bechara, 2005). Three models were compared. The first was the Expectancy Valence model (EV; Busemeyer & Stout, 2002). This model was originally created for studying individual differences in the IGT. It is a version of classical reinforcement learning models (such as the Delta learning model; Gluck & Bower, 1988; Rumelhart & McClelland, 1986; Sutton & Barto, 1998). The EV model, like other reinforcement learning models, assumes that sampling of the alternatives is based to the foremost extent on their relative outcomes. Each choice alternative is picked with a probability (or likelihood) determined by its expectancy (or its subjective value for the decision maker), which is a function of

its relative payoff. Other studies of autism have relied on this basic premise of reinforcement learning for cognitive models of categorization tasks (e.g., Kriete & Noelle, 2009) and attention tasks (e.g., Bjorne & Balkenius, 2005; Triesch, Jasso, & Deák, 2007).

We contrasted this assertion with two models assuming that sampling is based on the exploratory value of an alternative rather than its worth. These models follow the theoretical arguments stressing the value of exploration for reducing uncertainty in adaptive learning processes (see review in Erev & Gopher, 1999). Previous adaptive models have incorporated this idea by assigning a "reward" to exploratory behavior (Brafman & Tenenbholz, 2003; Gittins & Jones, 1974; Kaelbling, 1993; Wittmann, Daw, Seymour, & Dolan, 2008). We implemented this in the most parsimonious form by a model denoted as exploration-based adaptation (ExploBA), in which only exploration is rewarding to the decision maker. In this model, alternatives that are selected become less likely to be sampled in the future.

Additionally, a third model we studied rests on the assumption that the exploratory search is not totally unaffected by the value of the alternative, but rather takes place among alternatives deemed as adequate or appropriate (when an alternative is perceived as inadequate it is not explored anymore). This model implies that the sampling process is mostly based on exploration, but also responds to choice outcomes in a coarse manner (being closer to a "yes"/"no" decision). It was labeled as the ranged exploration-based adaptation (R-ExploBA) model.

The three models are fully described in the Appendix section. The ExploBA model includes only one parameter, a choice consistency parameter denoting the magnitude of the exploration based on the alternative's newness. When the value of this parameter is high, un-sampled alternatives are explored more; and when it is low past sampling plays a lesser role and choices are more random. The R-ExploBA model has an additional threshold parameter for determining when an alternative is inadequate.

The models were contrasted for their ability to predict one step ahead choices on each trial (for a review of this procedure, see Busemeyer & Diederich, 2009). The parameters of the models were optimized separately for each individual decision maker by maximizing the likelihood of the observed sequence of 100 choices produced by an individual. Optimization is a process wherein the fit of the model (in log likelihood) is compared with the fit of a baseline model. The baseline model's prediction is based on the optimized proportions of choices from the different decks. In the IGT the baseline model's three parameters are the average choice proportions of decks A, B, and C (deck D's is calculated accordingly). A comparison of the fit from the learning model to the baseline model is characterized by the improvement in the fit of the learning model over the baseline model. The log likelihood difference between the adaptive learning models and baseline model was corrected using the Bayesian Information Criterion (BIC; Schwartz, 1978). The BIC is a model-comparison index based on Bayesian principles which penalizes models with additional parameters:

$$\text{BIC} = 2 \cdot \log \text{likelihood difference} - k \cdot \ln(N) \quad (1)$$

Table 2

Model fit results: BIC scores for the Expectancy Valence (EV) model, the exploration-based adaptation model (ExploBA), and the ranged exploration-based adaptation (R-ExploBA). A comparison of the ASC and HC groups.

Group		EV	ExploBA	R-ExploBA
ASC	Median	0.23	-39.42	1.42
	Average	1.10	-52.34	-8.89
	SD	6.33	61.05	17.30
HC	Median	3.39	-59.24	-5.18
	Average	11.80	-78.49	-13.18
	SD	25.37	61.66	20.01

Table 3

Simplified Binary task results: mean proportions of selections from the advantageous alternative (P(A)) and average and largest run sizes for the ASC and HC groups.

Group		SBT1	SBT2	SBT3	SBT4	SBT5
ASC	P(A)	0.88	0.74	0.62	0.57	0.51
	Average run	4.37	3.40	5.71	3.78	4.44
	Largest run	15.67	14.33	15.11	13.55	7.89
HC	P(A)	0.85	0.83	0.64	0.52	0.54
	Average run	3.63	3.81	2.68	3.04	1.71
	Largest run	14.71	14.11	10.89	10.36	7.64

where k equals the difference between models in the number of parameters and N equals the number of observations (1 00). Positive values of the BIC statistic indicate that the cognitive model performs better than the baseline model.

The results of the analysis are summarized in Table 2. As can be seen, the fit for the ExploBA model was not adequate, being below that of the baseline model. The median fit of the EV model was adequate for the HC group (BIC = 3.39) but it was barely above the baseline model for the ASC group (BIC = 0.23). Under the R-ExploBA model the median fit in the HC group dropped (BIC = -5.18). However, for the ASC group it improved (to 1.42) compared to the EV model. An examination of individual decision makers shows that, compared to the EV model the R-ExploBA improved the accuracy of choice predictions for 50% of the participants in the ASC group, but only for 15% of the HC participants ($Z = 2.41, p < .01$). Accordingly, while not all participants in the ASC group appear to conform to the sampling process implied by the R-ExploBA model, it accounted for the choices of a substantial proportion of the participants in this group. The participants in the ASC group for whom the R-ExploBA fitted better than the EV had an average run of 1.1 and their largest run was 3.9. The participants for which the EV had better fit had an average run of 1.6 and a largest run of 10.7, and they were thus more similar to the control participants.

An examination of the best fits from both EV and R-ExploBA models shows that ASC participants were not much less predictable than healthy controls, with a combined median of 3.40 compared to 4.12 in controls ($Z = 0.72, p = .49$). This implies that the suggestion that ASC individuals simply have more error and are thus unpredictable has no basis. Note also that there were no significant differences between the ASC and HC groups in the EV or R-ExploBA model parameters. This indicates that what appears to be driving the differences between ASC adolescents and our control sample is not just one psychological parameter, such as the extent of the exploration. Rather, the complete adaptive learning mechanism appears to be different in a substantial proportion of ASC participants, with exploration replacing value as the main driver in the motivation to sample choice alternatives.

3.2. Behavioral implications of exploration-based adaptation

To illustrate the behavioral significance of the explorative adaptation pattern characterizing approximately 50% of the ASC sample, we analyzed the participants' responsiveness to losses, defined as the tendency to shift choices following a loss. Exploration-based learning would imply lower sensitivity to losses compared to no losses, as the contingent rewards have a smaller effect on behavior. The probability of changing one's choices following a loss (compared to no loss) in the two groups was calculated. For the HC group, the likelihood of switching choices after no losses was 56% and after losses it was 68%. This increase in switching following losses was significant ($t(27) = 3.36, p < .01$). For the ASC group, the likelihood of switching following no losses was 74% and following losses it was 73% ($t(14) = -0.16, p = .88$). Thus, the switching likelihood for ASC participants was almost equal following gain

and loss trials, indicating that they were unaffected by previous losses in their switching decisions. This finding is consistent with our analysis above, because exploration-based adaptation implies that outcomes matter less in sampling decisions.

3.3. Simplified Binary tasks

We next analyzed the results in the simplified binary decision tasks (see Table 3). A comparison of the maximization rate in these tasks revealed no significant group difference in the choice of the advantageous alternative, and no group by task-block interaction. There were also no significant differences between groups in the average and largest run. As can be seen in Table 3, these measures were quite similar in the two groups. This suggests that the different adaptation style of ASC participants emerges only in relatively difficult decision tasks.

3.4. Intellectual abilities

The average similarities score for the ASC group was close to that of the HC group (ASC: Average = 11.3, Median = 11.0; SD = 2.7; HC: Average = 11.1, Median = 11.5; SD = 2.7), with no significant difference between the two groups. However, on the non-verbal test of intellectual aptitude, the average score was lower for the ASC group (ASC: Average = 53.2, Median = 70; SD = 36.1; HC: Average = 82.4, Median = 90; SD = 19.1), with the difference between groups being statistically significant ($t(41) = 2.82, p < .01$). The relative strength in verbal ability and relative weakness in visual-spatial ability is common in Asperger's syndrome (Fitzgerald & Corvin, 2001), which constitutes the main clinical diagnosis in the ASC group.

4. General discussion

Individuals with ASC have a tendency to continuously shift between choice alternatives in complex repeated choice tasks (Johnson et al., 2006). In the current study, control participants made about 2.4 consecutive choices on each run, compared to 1.4 in the ASC sample. The ensuing pattern appears like pendulum movements from one choice alternative to the other, with choices being switched 73% of the time in the ASC group compared to 41% in the sample of healthy adolescents. An examination of datasets from various clinical populations on the Iowa Gambling task (Busemeyer & Diederich, 2009; Sevy et al., 2007; Yechiam et al., 2005, 2008) reveals that the extreme pattern of continuous switching is unique to ASC.³

Our study went beyond previous findings (e.g., Johnson et al., 2006; Minassian et al., 2006) in attempting to narrow down the range of tasks in which this pattern appears. Our examination of

³ The only exception appears to be neurological patients with right somatosensory and insular damage, who show a similar pattern (Yechiam et al., 2005). However, an examination of this population's dataset with the current exploration-based models reveals that unlike in ASC, their shifting choice pattern is driven mostly by random error.

the various Two-Button tasks showed that in tasks where the advantageous alternatives are easy to identify, there were no significant group differences. It appears that the increased sampling is found only in relatively complex tasks. Our modeling analysis sheds light on the reasons for that.

While many of the performance deficits in ASC indeed appear to emerge in complex tasks (Minshew & Goldstein, 1998), there may be different explanations for why this would be the case in repeated choice tasks. We examined three possible explanations: an explanation based on an increase in error due to cognitive difficulties, an explanation based on a quantitative increase in exploration, and an explanation based on a qualitative increase in exploration such that it becomes the main directive of the adaptation process. Our analysis was most consistent with the third explanation. According to our modeling results, for a substantial proportion of ASC individuals facing the complex decision task, the explorative value of choice alternatives (i.e., the extent to which they were previously un-explored) took precedence over their outcome value. Outcomes were not neglected completely, but they were integrated in a coarser fashion, with alternatives producing very poor outcomes being taken out from the range of select-able alternatives. This suggests that what leads to the pattern of shifting choices in the IGT is an extensive drive to explore the un-sampled alternatives.

More specifically, while a conventional reinforcement learning model (the EV model for the IGT) had good fit to the control adolescents data, it did not predict choices in the ASC group with great accuracy (as in Johnson et al., 2006). Yet this was not due to the behavior of the ASC group being simply more random. A modified model based on exploration as the main directive of the adaptation process and outcome value as a secondary directive (R-ExploBA) was successful in improving the average predictions for the ASC group beyond a baseline statistical model. It should be emphasized that this pattern was not observed for all of the participants in the ASC group. As in previous studies (e.g., Johnson et al., 2006) large individual differences were found in this group, with the exploration chasing mechanism taking precedence in about half of the participants (compared to 15% in the control group).

Our results thus suggest that some individuals with ASC have a form of adaptation characterized by being less sensitive to the immediate incentive structure and by an intensive exploratory search of the available alternatives. Our findings further suggest that this kind of exploration-based behavior is triggered in relatively complex choice tasks, and is not applied uncontrollably in tasks that have a clear favorable solution (or clear rejects). We use the term “exploration-based adaptation” rather than “novelty based” because the stimuli were not novel in the sense of being striking or surprising; they just contained a pattern that was uneasy to learn. Our terminology is consistent with the findings that individuals with ASC have no increased novelty seeking on the Temperament and Character Inventory (TCI; Cloninger, Przybeck, Svrakic, & Wetzell, 1994), although they do show low reward dependence, which denotes heightened self-directedness (e.g., Anckarsäter et al., 2006; Sizoo, van den Brink, van Eenige, & van der Gaag, 2009; Soderstrom, Rastam, & Gillberg, 2002). Additionally, various studies have found that individuals with ASC are more distracted by novelty than healthy individuals (e.g., Sokhadze et al., 2009), and have trouble processing novel stimuli (e.g., Gomot et al., 2006). These previous studies suggest that the exploratory adaptation pattern of people with ASC represents a drive towards solving a problem, rather than towards exciting novel occurrences.

Although we present a small group of patients, we believe that our results actually reflect the general pattern in ASC. One possible implication of this finding is that ASC may have an adaptive role in encouraging people, especially gifted individuals, to persevere despite of what appears to be failure. Previous studies of autism have suggested that children and adolescents with ASC are less creative

and imaginative (e.g., Craig & Baron-Cohen, 1999), yet these studies have relied for the most part on tests of divergent thinking, which assess high level abstractization and plasticity of thought (such as the Torrance test battery; Torrance, 1974). However, creativity is a complex process which also relies on convergent thinking, referring to the ability to focus on a problem and explore its different sides and aspects (Guilford, 1967; see also Fung, 2009). In such convergent processes, the explorative style of ASC individuals may offer an advantage. In particular, enhanced exploration is beneficial for solving problems characterized by having multiple possible solutions with some of the solutions being locally optimal (Fu & Gray, 2004).

It should be noted though that the same profile of intensive exploration can lead to adaptive failure in situations where the boundaries for safe exploration are narrow and the environment is relatively unforgiving. For example, Mackinlay, Charman, and Karmiloff-Smith (2006) examined the performance of high functioning children with ASC and controls in a set of multiple tasks with strict conditions about the order of switching between tasks. The results showed poor performance on the part of ASC individuals as well as more instances of inefficient switching between tasks (where the rules of the game were broken).

The current findings may have relevant implications to the well-known gap between academic and social performance in individuals with ASC (see Klin, Jones, Schultz, & Volkmar, 2003; Macintosh & Dissanayake, 2006). Inappropriate social behaviors and social failures have been attributed to difficulties in interpreting emotional cues (e.g., Baron-Cohen, 2009; Wang, Lee, Sigman, & Dapretto, 2006), and to learning impairments in a task requiring the integration of multiple signals of salience and reward (Klinger et al., 2006; Minshew & Goldstein, 1998). However, social difficulties could also be due to differences in adaptation style. Klin et al. (2003) proposed that the motivation to orient to salient social stimuli and to seek social meaning may be underlie the social difficulties in ASC (for supporting findings see Baron-Cohen, Baldwin, & Crowson, 1997; Klin, 2000). Our findings demonstrate a specific mechanism of over-exploration which distinguishes the adaptation pattern of ASC individuals from that of healthy controls. This mechanism, involving detailed exploration of the various alternatives, may be inappropriate in a social environment characterized by dynamic and fluent changes and the need for speeded responses. Previous studies that have used decision tasks to predict socialization in ASC focused on the skill of reward learning, and did not examine adaptation style (e.g., Dawson et al., 2002; Munson, Faja, Meltzoff, Abbott, & Dawson, 2008). This is an interesting and important direction for future studies.

Acknowledgment

This research was supported in part by the Israel Science Foundation (Grant No. 244/06) and by the Max Wertheimer Minerva Center for Cognitive Studies.

Appendix A

A.1. The Expectancy Valence (EV) model

This reinforcement learning model is composed of three basic components:

- (1) The evaluation of the gains and losses experienced after making a choice is called a valence. The valence is denoted $u(t)$, and is calculated as a weighted average of gains and losses from the chosen option in trial t .

$$u(t) = w \cdot \text{win}(t) - (1 - w) \cdot \text{loss}(t) \quad (1a)$$

where $win(t)$ and $loss(t)$ are the amounts of money won or lost on trial t ; and w is the weight parameter indicating the subjective weight to gains versus losses ($0 \leq w \leq 1$).

- (2) The valences produced by a deck j are summarized by an accumulated subjective value for each deck, called an expectancy, and denoted $E_j(t)$. A Delta learning rule (Estes & Burke, 1953) is used for updating the expectancy after each choice:

$$E_j(t) = E_j(t-1) + \phi \cdot [u(t) - E_j(t-1)] \quad (2a)$$

where j is the selected deck. The recency parameter, ϕ , describes the degree to which expectations about deck consequences reflect the influence of the most recent outcomes or more distant past experience ($0 \leq \phi \leq 1$).

- (3) The probability of choosing a deck is a strength ratio of the expectancy of that deck relative to all decks, using Luce's (1959) rule:

$$\Pr[G_j(t+1)] = \frac{e^{\theta(t) \cdot E_j(t)}}{\sum_j e^{\theta(t) \cdot E_j(t)}} \quad (3a)$$

where $\Pr[G_j(t)]$ is defined as the probability that deck j will be selected on trial t by the model, and where: $\theta(t) = (t/10)^c$. Here c is the choice consistency parameter ($-5 \leq c \leq 5$) controlling the consistency of the choice probabilities and the expectancies.

A.2. The exploration-based adaptation (ExploBA) model

In this model choices are not driven by expectancies. Rather, they are driven by the exploratory value of the alternative, defined by the relative past samples from the alternative. The exploratory value is calculated as follows:

$$V_j(t) = t/k - N_j(t) \quad (4a)$$

where $V_j(t)$ denotes the exploratory value of each alternative, k denotes the number of alternatives, and $N_j(t)$ denotes the number of selections from alternative j by time t . The sampling therefore conforms to a simple rule ensuring that the more an alternative is chosen the less it is likely to be sampled in the future. The exploratory values are in turn inserted into formula (3a) instead of the expectancies ($V_j(t)$ replaces $E_j(t)$ in this equation).

A.3. The ranged exploration-based adaptation (R-ExploBA) model

In this model, exploration takes place within a certain range of alternatives (see Ratcliff, 1978 for a similar idea in memory models). The following formula is added to the ExploBA model for determining this range:

$$\text{If } \frac{e^{E_j(t)}}{\sum_j e^{E_j(t)}} < \varepsilon, \text{ then alternative } j \text{ becomes erased} \quad (5a)$$

Parameter ε is the threshold parameter defining the inadequacy of a choice alternative ($0 \leq \varepsilon \leq 1$). An erased alternative that falls below the threshold is no longer selected. For parsimony, the expectancies are simply the last outcomes produced by a choice alternative (when a severe enough outcome is encountered, the alternative is no longer sampled). Having the expectancies as the mean experienced outcomes yields almost identical results.

References

Ånckarsäter, H., Stahlberg, O., Larson, T., Hakansson, C., Jutblad, S. B., Niklasson, L., et al. (2006). The impact of ADHD and autism spectrum disorders on

- temperament, character, and personality development. *American Journal of Psychiatry*, 163, 1239–1244.
- Ashburner, J., Ziviani, J., & Rodger, S. (2009). Surviving in the mainstream: Capacity of children with autism spectrum disorders to perform academically and regulate their emotions and behavior at school. *Research in Autism Spectrum Disorders*. doi:10.1016/j.rasd.2009.07.002.
- Baird, G., Charman, T., Cox, A., Baron-Cohen, S., Swettenham, J., Wheelwright, S., et al. (2001). Screening and surveillance for autism and pervasive developmental disorders. *Archives of Disease in Childhood*, 84, 468–475.
- Baron-Cohen, S. (2009). Autism: The empathizing–systemizing (E–S) theory. *Annals of the New York Academy of Sciences*, 1156, 68–80.
- Baron-Cohen, S., Baldwin, D. A., & Crowson, M. (1997). Do children with autism use the speaker's direction of gaze strategy to crack the code of language? *Child Development*, 68, 48–57.
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism Spectrum Quotient (AQ): Evidence from Asperger syndrome/high functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, 31, 5–17.
- Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, 50, 7–15.
- Bjorne, P., & Balkenius, C. (2005). A model of attentional impairments in autism: First steps toward a computational theory. *Cognitive Systems Research*, 6, 193–204.
- Brafman, R. I., & Tenenbholz, M. (2003). R-MAX – A general polynomial time algorithm for near-optimal reinforcement learning. *Journal of Machine Learning Research*, 3, 213–231.
- Busmeyer, J. R., & Diederich, A. (2009). *Cognitive modeling*. Thousand Oaks, CA: Sage.
- Busmeyer, J. R., & Stout, J. C. (2002). A contribution of cognitive decision models to clinical assessment: Decomposing performance on the Bechara gambling task. *Psychological Assessment*, 14, 253–262.
- Cloninger, C. R., Przybeck, T. R., Svrakic, D. M., & Wetzel, R. D. (1994). *The temperament and character inventory (TCI): A guide to its development and use*. St. Louis, MO: Center for Psychobiology of Personality, Washington University.
- Craig, J., & Baron-Cohen, S. (1999). Creativity and imagination in autism and Asperger's syndrome. *Journal of Autism and Developmental Disorders*, 29, 319–326.
- Dawson, G., Munson, J., Estes, A., Osterling, J., McPartland, J., Toth, K., et al. (2002). Neurocognitive function and joint attention ability in young children with autism spectrum disorder versus developmental delay. *Child Development*, 73, 345–358.
- Erev, I., & Gopher, D. (1999). Cognitive game-theoretic analysis of attention strategies, ability and incentives. In D. Gopher & A. Koriati (Eds.), *Attention and performance, XVIII: Cognitive regulation of performance, theory and applications*. Cambridge: MIT Press.
- Estes, W. K., & Burke, C. J. (1953). A theory of stimulus variability in learning. *Psychological Review*, 6, 276–286.
- Fitzgerald, M. (2004). *Autism and creativity: Is there a link between autism in men and exceptional ability?* East Sussex, UK: Brunner-Routledge.
- Fitzgerald, M., & Corvin, A. (2001). Diagnosis and differential diagnosis of Asperger syndrome. *Advances in Psychiatric Treatment*, 7, 310–318.
- Fu, W.-T., & Gray, W. D. (2004). Resolving the paradox of the active user: Stable suboptimal performance in interactive tasks. *Cognitive Science*, 28, 901–935.
- Fung, C. H. M. (2009). Asperger's and musical creativity: The case of Erik Satie. *Personality and Individual Differences*, 46, 775–783.
- Gillberg, C. (2002). *A guide to Asperger's syndrome*. Cambridge: University Press.
- Gittins, J. C., & Jones, D. M. (1974). A dynamic allocation index for the sequential design of experiments. In J. Gani (Ed.), *Progress in statistics* (pp. 241–266). North-Holland: Amsterdam.
- Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General*, 118, 309–331.
- Golan, O., Baron-Cohen, S., Hill, J. J., & Rutherford, M. D. (2007). The 'reading the mind in the voice' test-revised: A study of complex emotion recognition in adults with and without autism spectrum conditions. *Journal of Autism and Developmental Disorders*, 37, 1096–1106.
- Gomot, M., Bernard, F. A., Davis, M. H., Belmonte, M. K., Ashwin, E. T., Bullmore, E. T., et al. (2006). Change detection in children with autism: An auditory event-related fMRI study. *Neuroimage*, 29, 475–484.
- Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill Book Company.
- Johnson, S. A., Yechiam, E., Murphy, R. M., Queller, S., & Stout, J. C. (2006). Motivational processes and autonomic responsivity in Asperger's disorder: Evidence from the Iowa gambling task. *Journal of the International Neuropsychological Society*, 12, 668–676.
- Kaelbling, L. P. (1993). *Learning in embedded systems*. Cambridge, MA: MIT Press.
- Klin, A. (2000). Attributing social meaning to ambiguous visual stimuli in higher-functioning autism and Asperger syndrome: The social attribution task. *Journal of Child Psychology and Psychiatry*, 41, 831–846.
- Klin, A., Jones, W., Schultz, R., & Volkmar, F. (2003). The enactive mind, or from actions to cognition: Lessons from autism. *Philosophical Transactions of the Royal Society London B*, 358, 345–360.
- Klinger, L. G., & Dawson, G. (2001). Prototype formation in autism. *Development and Psychopathology*, 1, 111–124.

- Klinger, L. G., Klinger, M. R., & Pohlig, R. A. (2006). Implicit learning impairments in autism spectrum disorders: Implications for treatment. In J. M. Perez, P. M. Gonzalez, M. L. Comi, & C. Nieto (Eds.), *New developments in autism*. London: Jessica Kingsley.
- Kobayashi, R., & Murata, T. (1998). Behavioral characteristics of 187 young adults with autism. *Psychiatry and Clinical Neurosciences*, 52, 383–390.
- Kriete, T.E., & Noelle, D.C. (2009). Impaired cognitive flexibility and intact cognitive control in autism: A computational cognitive neuroscience approach. In *Proceedings of the Annual Conference of the Cognitive Science*, Stresa, Italy.
- Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 24, 659–686.
- Luce, R. D. (1959). *Individual choice behavior*. New York, NY: Wiley.
- Macintosh, K., & Dissanayake, C. (2006). Social skills and problem behaviors in school aged children with high-functioning autism and Asperger's disorder. *Journal of Autism and Developmental Disorders*, 36, 1065–1076.
- Mackinlay, T., Charman, T., & Karmiloff-Smith, A. (2006). High functioning children with autism spectrum disorder: A novel test of multitasking. *Brain and Cognition*, 61, 14–24.
- Minassian, A., Paulus, M., Lincoln, A., & Perry, W. (2006). Adults with autism show increased sensitivity to outcomes at low error rates during decision-making. *Journal of Autism and Developmental Disorders*, 37, 1279–1288.
- Minshew, N. J., & Goldstein, G. (1998). Autism as a disorder of complex information processing. *Mental Retardation and Developmental Disabilities Research Reviews*, 4, 129–136.
- Munson, J., Faja, S., Meltzoff, A., Abbott, R., & Dawson, G. (2008). Neurocognitive predictors of social and communicative developmental trajectories in preschoolers with autism spectrum disorders. *Journal of the International Neuropsychological Society*, 14, 956–966.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59–108.
- Raven, J. C. (1965). *Progressive matrices*. New York: The Psychological Corporation.
- Rumelhart, D. E., & McClelland, J. E. The PDP Research Group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vols. 1 and 2). Cambridge, MA: MIT Press.
- Schwartz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 5, 461–464.
- Sevy, S., Burdick, K., Visweraiah, H., Abdelmassih, S., Lukin, M., Yechiam, E., et al. (2007). Iowa Gambling Task in schizophrenia: A review and new data in patients with schizophrenia and co-occurring cannabis use disorders. *Schizophrenia Research*, 92, 74–84.
- Sizoo, B., van den Brink, W., van Eenige, M. G., & van der Gaag, R. J. (2009). Personality characteristics of adults with autism spectrum disorders or attention deficit hyperactivity disorder with and without substance use disorders. *Journal of Nervous and Mental Disease*, 197, 450–454.
- Soderstrom, H., Rastam, M., & Gillberg, C. (2002). Temperament and character in adults with Asperger syndrome. *Autism*, 6, 287–297.
- Sokhadze, E., Baruth, J., Tasman, A., Sears, L., Mathai, G., El-Baz, A., et al. (2009). Event-related potential study of novelty processing abnormalities in autism. *Applied Psychophysiology and Biofeedback*, 34, 37–51.
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.
- Torrance, E. P. (1974). *Norms technical manual: Torrance tests of creative thinking*. Lexington, Mass: Ginn and Co.
- Triesch, J., Jasso, H., & Deák, G. O. (2007). Emergence of mirror neurons in a model of gaze following. *Adaptive Behavior*, 15, 149–165.
- Wang, T., Lee, S. S., Sigman, M., & Dapretto, M. (2006). Neural basis of irony comprehension in children with autism: The role of prosody and context. *Brain*, 129, 932–943.
- Wechsler, D. (1981). *Manual for the wechsler adult intelligence scale (rev.)*. New York: Psychological Corporation.
- Wittmann, B. C., Daw, N. D., Seymour, B., & Dolan, R. J. (2008). Striatal activity underlies novelty-based choice in humans. *Neuron*, 26, 967–973.
- Yechiam, E., Busemeyer, J. R., Stout, J. C., & Bechara, A. (2005). Using cognitive models to map relations between neuropsychological disorders and human decision making deficits. *Psychological Science*, 16, 973–978.
- Yechiam, E., & Ert, E. (2007). Evaluating the reliance on past choices in adaptive learning models. *Journal of Mathematical Psychology*, 51, 75–84.
- Yechiam, E., Kanz, J., Bechara, A., Stout, J. C., Busemeyer, J. R., Altmaier, E. M., et al. (2008). Neurocognitive deficits related to poor decision-making in people behind bars. *Psychonomic Bulletin and Review*, 15, 44–51.