Development and Psychopathology, **13** (2001), 97–110
Copyright © 2001 Cambridge University Press
Printed in the United States of America

The execution of arbitrary procedures by children with autism

SZILVIA BÍRÓ AND JAMES RUSSELL
University of Cambridge

Abstract
Executive tasks typically contain a prepotent lure. In addition, they require individuals either to (a) follow arbitrary procedures or (b) update their model of the physical world. Recent research suggests that children with autism may be challenged only by executive tasks of the former kind (containing arbitrary rules). We asked whether this continues to be true when there is *no* prepotent lure, comparing performance on (a) a task with arbitrariness but without prepotency with (b) a task with both features and (c) a task with neither. The participants with autism performed at a lower level than comparison groups on the first and second task but not on the third task. This outcome is consistent with the view that autism is associated with difficulties in following arbitrary procedures per se. We try to locate difficulties with acting on the basis arbitrary rules in relation to more mainstream ideas about autistic executive dysfunction.

While the sensorimotor development of children with autism appears to be normal (Abrahamsen & Mitchell, 1990; Adrien, Marineau, Barthélémy, Bruneau, Garreau, & Savage, 1995), their performance on formal tests of executive functioning is generally found to be impaired (Ozonoff, Pennington, & Rogers, 1991; chapters in Russell, 1997a). The term "executive functions" refers to a group of domain-general cognitive processes responsible for planning and guiding complex behaviors: the instigation of appropriate actions, disengaging from strategies and engaging new ones, inhibiting prepotent but incorrect responses, monitoring the effects of actions, representing future actions and action-relevant information.

Since the initial studies reporting executive impairment in autism (Rumsey, 1985; Rumsey & Hamburger, 1988), research has been aimed at specifying which among the many kinds of executive function are affected. This task is an urgent one for two reasons. First, executive impairments are also present in a variety of other developmental disorders, such as attention-deficit hyperactivity disorder, Tourette's syndrome, obsessive–compulsive disorder, and early-treated phenylketonuria (Ozonoff, 1997, for a review). Given this, an executive account of the core cognitive dysfunctions in autism can only be convincing if a qualitatively different profile of executive impairments is demonstrated within the disorder. Second, specifying the nature of the executive deficits in autism will help to clarify the relation between the executive deficits and the other well-research deficits in autism such as those in "theory of mind" (e.g., Baron–Co-

This work was carried out by the first author in partial fulfilment of the requirements of the degree of MPhil at Cambridge University, and under the supervision of the second author. We are very grateful to the staff and pupils of the following schools for their kind cooperation: Harborough School, London; Rosehill School Nottingham; Wargrave House School, Newton-le-Willows; Priory School, Cambridge; St. Luke's School, Cambridge; Marshfield School, Peterborough; Rees Thomas School, Cambridge.

Address correspondence and reprint requests to: James Russell, Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK; E-mail: jr111@cus.cam.ac.uk.

https://doi.org/10.1017/S0954579401001079 Published online by Cambridge University Press

hen, Leslie & Frith, 1985) and in "central coherence" (Frith & Happé, 1994). We will be reporting a study of means–end behavior in children with autism that aims to address such questions about the locus of their executive disorders.

While the tests on which executive functioning is assessed vary in their surface form, two underlying demands are typically present: the need to (a) hold in mind task-relevant information while (b) inhibiting prepotent responses.[1] These demands are certainly present in classic "frontal" tasks like the Wisconsin Card Sorting Task (WCST) and its cognate task the Intra-Dimension/Extra-Dimension Shift task, as well as in the Tower of Hanoi (TOH) and Tower of London (TOL) planning tasks. In WCST the participant has to engage a new, arbitrary sorting rule after receiving negative feedback and do so while the prepotent response to continue sorting by a previously correct category is being inhibited. The TOH task requires the participant to hold in mind a sequence of future moves together with certain rules of procedure (e.g., larger discs cannot be placed above smaller). In this case the prepotent response that the participant must inhibit is putting the disk or ball directly onto the target peg without acknowledging the need to make detour moves. Several studies have demonstrated a robust impairment on such tasks in children with autism (e.g., Ozonoff, Pennington, & Rogers, 1991; Hughes, Russell, & Robbins, 1994).

With particular regard to the development of means–end behavior, Hughes and Russell (1993) designed a task (the "switch task") intended to gauge children's difficulties with rule following plus response inhibition in the context of goal-directed action. This task requires the participant to retrieve a marble resting on a plinth inside a box by reaching through a hole in the box's front face. A direct reach (the prepotent but wrong response) will break an invisible light-beam at the entrance to the box which in turn trips a mechanism causing the marble to drop down inside the plinth. To prevent this happening participants have to hold in mind that it is necessary to throw a switch on the side of the box before reaching. At the outset, the experimenter demonstrates this procedure. Successful performance was taken to be a run of three switch–reach sequences with 15 trials. Hughes and Russell (1993, Experiment 2) reported that this task revealed a degree of impairment in autism as least as great as that found on theory-of-mind tasks such as false belief (Baron–Cohen, et al., 1985). Failure almost invariably took the form of reaching directly. While normally developing 3-year-olds and verbal–mental age (VMA) matched children with mental handicap were rarely challenged by the task, participants with autism often failed it. Some of these had VMAs as high as 10 years.

What motivated the present study, however, was the fact that there exist tasks that are executive on the above definition on which children with autism are *not* impaired. Some of these are Piagetian search tasks. As Diamond (1985, 1988) argued, these are executive in the sense that they require the child to carry out a goal-directed action based on action-relevant information despite a conflicting habitual tendency. In the A-not-B task, for example, children retrieve a desired object a number of times from behind occluder A and then witness its transfer to occluder B. In Diamond's terms, they must therefore inhibit the prepotent response of reaching back to occluder A and maintain the information that the location of the object changed from A to B. Neither McEvoy, Rogers and Pennington, (1993), testing preschool children with autism, nor Wehner and Rogers (1994), nor Pennington, Rogers, Bennetto, Griffith, Reed, and Shyu (1997, preliminary results), found children with autism to be impaired on this task.

Another object-retrieval task with an executive character is the "tubes" task (Hood, 1995). On this task, too, children with autism are not impaired (Russell, Jarrold, & Hood, 1999, Experiment 1). Moreover, like Hughes and Russell's switch task, the tubes task is normally failed below the age of about 3

1. "Prepotent responses" can be made when there is possibility of making an erroneous response that is called out either by some salient feature of the environment or by some feature rendered salient through previous learning.

https://doi.org/10.1017/S0954579401001079 Published online by Cambridge University Press

years. In it the child, being faced with an opaque tube that curves diagonally downwards to one of three catch trays, is required to visit the appropriate catch tray when a ball is dropped into the opening of the tube. The dominant error is to search in the tray that is immediately beneath the mouth of the tube *despite the fact that this tray is not connected to the tube*. It outlets into the adjacent tray. Here too the task encourages a prepotent response, in the sense of a response based on the natural assumption that objects fall vertically. Whether one regards such vertical search as an overlearned procedure, or as manifesting innately specified "folk physics," or in some other way, it is clearly an "unthinking," habitual, but wrong, response—and accordingly prepotent. The error is very unlikely to have been due to a *conceptual* failure to appreciate that objects can move when not perceived, because we know that this knowledge is present at 18 months (Piaget, 1954), if not long before (e.g., Baillargeon, 1993; Ahmed & Ruffman, 1998).

Why then do children with autism perform poorly on certain executive tasks while showing no impairment on others? Russell (1997b) suggests that the divergent performance can be explained in terms of the nature of the information that has to be held in mind. In tasks such as WCST, TOH, and the switch task, the information that is in competition with the prepotent response is in the form of a novel and arbitrary rule—"arbitrary" in the sense of lacking a rationale. Recall that in WCST this rule is "sort by one dimension until you get negative feedback, then change to another," and in TOH it is that "larger discs cannot be placed above smaller." Similarly in the switch task, although the act of throwing a switch is familiar to children, switching in order to enable a retrieval act is certainly novel. With regard to arbitrariness, throwing the switch before reaching is an arbitrary procedure in the sense that there is nothing in the immediate environment to tell the child what has to be done. Moreover, the act has no rationale unless the participants happen to know that there is a light beam, what it effects, and the role of switching in preventing it. In contrast, in the other kinds of executive tasks (e.g., A-not-B, tubes task) the information that has to be held in mind concerns new states of the world—not a new arbitrary rule. There is also a rationale for basing their action on this information, which is one based on knowledge of folk physics.

As a way to test this conjecture, we asked whether children with autism have difficulty following arbitrary procedures per se. That is to say, we asked whether they have difficulty in completing a means–end task in which the means is arbitrary but in which there is *no prepotent lure*.

Participants were presented with three means–end tasks: one containing neither a prepotent lure nor an arbitrary procedure (–P–A), one with both (+P+A), and one (the crucial condition) with no prepotency but with an arbitrary procedure (–P+A). The tasks required the children to perform the same act—namely, retrieving a ball from a box by reaching through a hole. The first means–end task was the "lever task" (–P–A). In this, children had to depress a lever on the side of the box in order to release an opaque screen covering the hole on the box, after which they could reach through the hole for the ball. Note that here the means action was not arbitrary in relation to the goal as there was an intelligible physical relationship between the two, and note also that it was not possible to make a direct reach for the object, which means that there was no prepotent lure.

The second means–end task was the "switch task" (+P+A). This task was similar to the switch task of Hughes and Russell (1993), with the only difference being that in our task children did not have to monitor lights on top of the box telling them whether switching or an alternative retrieval procedure was required.

The third task was called the "cup task" (–P+A). The aperture in the box was blocked by an opaque screen that the child could release by inverting a cup located in front of the box. The cup *was not connected to the box*. In fact it was the experimenter who surreptitiously released the screen by remote control when the cup had been inverted. No prepotent response was possible in this case insofar as a direct reach was blocked by the opaque

https://doi.org/10.1017/S0954579401001079 Published online by Cambridge University Press

screen. Note that the means act fulfilled a criterion for being an arbitrary procedure insofar as cup inversion possesses no rationale in relation to the goal. There was no intelligible physical relationship between the cup inversion and the retrieval of the ball.

The central question is, then, whether children with autism will be impaired on the cup task (–P+A). If they have specific difficulties with following arbitrary rules (i.e., even when there is no prepotent lure present), then they should be impaired on that task as well as on the switch task (+P+A). If they have difficulties both with prepotency and with following an arbitrary rule, then they should perform worse on the switch task (+P+A) than on the cup task (–P+A). If they are only challenged when they need to combine inhibiting a prepotent response with following an arbitrary procedure, they will be impaired only on the switch task (+P+A). The performance of children with autism was compared against that of matched groups of normally developing children and that of children with moderate learning difficulties.

Method

Participants

Thirty children and adolescents with autism (AUT) participated in this study. They had all received a formal diagnosis of autism according to established criteria (American Psychiatric Association, 1987) and were attending special schools in London, Nottingham, or Merseyside. They had a mean chronological age of 11 years (range = 8 years 1 month to 19 years 4 months; *SD* = 3.29 years). There were also 30 children in the group with moderate learning difficulties (MLD). These were attending special schools in Cambridge and Peterborough. They had a mean chronological age of 12 years and 11 months (range = 4 years 11 months to 16 years 11 months; *SD* = 3.13 years). Thirty children from mainstream schools (NOR) in Cambridge also took part. This group had a mean age of 6 years 3 months (range = 2 years 10 months to 8 years 8 months; *SD* = 1.32 years).

Mental age assessment

VMA was assessed by the long form of the British Picture Vocabulary Scale (BPVS; Dunn, Dunn, Whetton, & Pintilie, 1982). The mean VMA for the group with autism and the group with moderate learning difficulties was 5 years 8 months (range = 2 years 4 months to 11 years 1 month; *SD* = 2.2 years, and range = 2 years 10 months to 10 years 10 months; *SD* = 2.15 years, respectively). The mean VMA for normally developing children was 5 years 7 months (range = 2 years 4 months to 11 years 6 months; *SD* = 2.25 years). Verbal IQ scores were also calculated for the two clinical groups: mean verbal IQ for AUT group was 46.22 (range = 16.05–82.47, *SD* = 20.22), mean verbal IQ for MLD group was 54.02 (range 17.73–92.85, *SD* = 17.89).

There were no significant differences in VMA between the AUT and MLD groups, t (58) = –0.13, p < 0.894, or between the AUT and normally developing groups, t (58) = 0.03, $p < 0.977$. A one-way analysis of variance (ANOVA) revealed no significant difference in VMA for the three order conditions (see Design section) in any group, F (2, 87) = 0.214, $p < 0.808$. Although children with autism were somewhat older than the children with MLD, there was no significant difference in verbal IQ between the AUT and MLD groups, t (58) = –1.58, $p < 0.119$.

Tasks and materials

Lever task (–P–A). A 30 × 30 × 30 cm wooden box was employed. There was a centrally located, circular (15 cm in diameter) opening in its front face. A cylindrical pedestal was placed behind the opening upon which an orange, bouncy ball rested. There was a 3-cm-long metal lever and a 1.5-cm diameter metal knob on the left side of the box. Turning the knob rotated a curved opaque screen which locked in between the opening and the ball, preventing the retrieval of the ball. When the screen was in "blocking" position, the ball was not visible. By pressing the lever the screen was released and sprang back, allowing access to the ball. A small door (15 × 20

https://doi.org/10.1017/S0954579401001079 Published online by Cambridge University Press

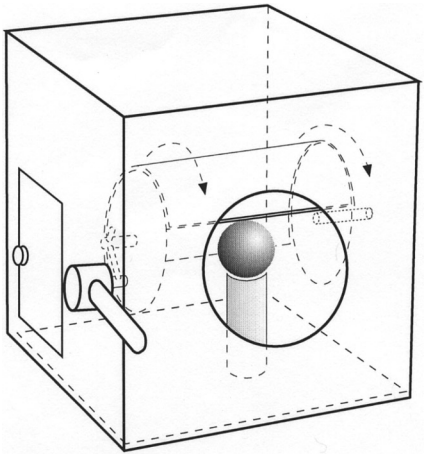


Figure 1. The box used in the "lever task."

cm) on the left side of the box was used to reposition the ball (see Figure 1).

Cup task (–P+A). The same box was used as in the lever task but with two modifications. Access to the lever was blocked by attaching a wooden box ($7 \times 7 \times 30$ cm) to the left-hand corner of the box (see Figure 2). The participants could neither see nor could reach the lever. The knob functioned in the same way as in the lever task: it rotated the screen to its blocking position. Unseen by the participant, the experimenter manipulated a piece of apparatus of the following kind. A plastic tube ending in a metal pin was connected to the screen, and this tube passed through a small hole on the back side of the box. There was a small rubber bulb on the other end of the tube which functioned as a pump to push the pin on the other end of the cable, thus releasing the screen that sprang back, allowing access to the ball. A cup was placed on the table approximately 10 cm from the front of the box. It was made clear to the participants that the cup was not physically connected to the box. As soon as the cup was inverted, the experimenter surreptitiously operated the screen releaser by squeezing the bulb, thus creating the impression that turning over the cup caused the screen to spring back.

Switch task (+P+A). An aluminum box was used in this task which was also $30 \times 30 \times 30$ cm (see Figure 3). There was a centrally

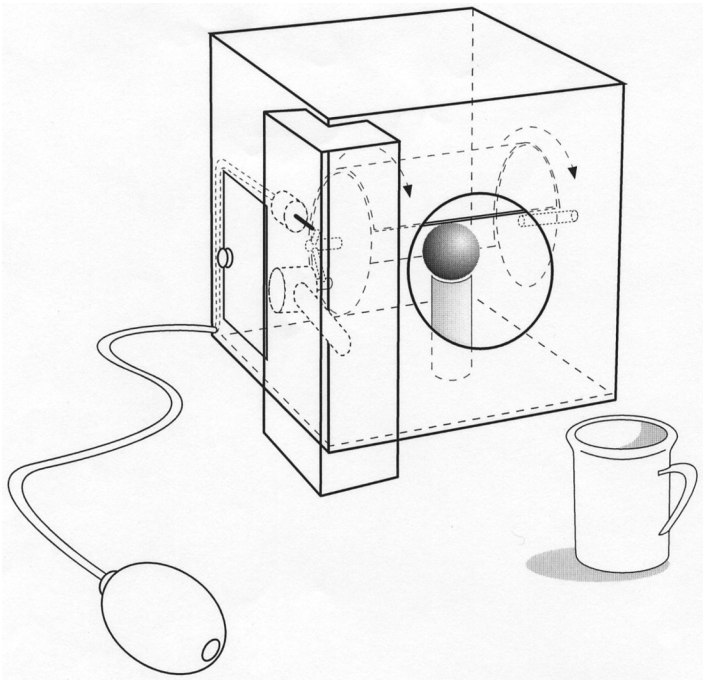


Figure 2. The box used in the "cup task."

https://doi.org/10.1017/S0954579401001079 Published online by Cambridge University Press

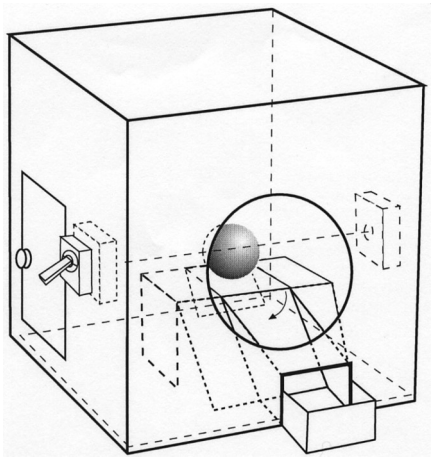


Figure 3. The box used in the "switch task."

located, circular opening (15 cm in diameter) on its front face, cut into a transparent Perspex square. Mounted inside the box was a pyramid-shaped platform with a narrow, flat top on which a small rubber ball rested. On both sides of the opening, photoelectric cells emitted infrared beams directed towards the top of the platform. When the beams were broken, a trapdoor at the top of the platform was automatically activated causing the ball to drop out of sight. On the left-hand side of the box there was a 2-cm-long metal switch. Operating the switch deactivated the light-beam trigger and thereby allowed access to the ball. There was a further small door (15 × 20 cm) on the left side of the box through which the experimenter could reposition the ball at the end of each trial. (Figure 3 also shows that there was a chute running down from the aperture. In Hughes & Russell's study, 1993, this had provided an alternative way of retrieving the ball from the box. Note that this part of the apparatus was not used in the present task and did not interfere with performance in any way.)

A Sony Video 8 Handicam was used to record the children's performance for later data analysis.

Design

Each participant in each group (AUT, MLD, NOR) was tested on all three tasks. The tasks were presented in three different orders. Equal numbers of children in each group received the tasks in lever–cup–switch, cup–switch–lever, or switch–lever–cup orders. Accordingly there were 10 children in each order group. Thus, the within-subject variable of task was supplemented by two orthogonal between-subject factors: group and task order.

Procedure

Participants were tested individually in a quiet room in their school in one session that lasted about 15 min. First, an orange, bouncy ball was given to the children, and they were allowed to play with it for a short time. The purpose of this warming-up period was to generate interest and build up some motivation to retrieve the ball during the testing session.

In the lever–cup–switch condition the wooden box was introduced initially. The ball was placed on the pedestal inside the box and the participant was asked to retrieve it twice. Then the screen was lowered and the participant was again invited to try to take the ball out. After a few unsuccessful attempts, participant and the experimenter agreed that they could not retrieve the ball when the screen was down. Next, the experimenter drew the participant's attention to the lever. She then showed the participant, with accompanying explanation, that the only way to obtain the ball was to press the lever that released the screen. The experimenter demonstrated this two times and then asked the participant to do the same. Each time there was failure to retrieve the ball (see Scoring section below) the correct procedure was demonstrated and also explained. The criterion for the successful completion of each task was three consecutive and coordinated lever-and-reach sequences within 15 trials. When the criterion had been reached, or 15 trials had been attempted (which ever came first), the next task was introduced.

The experimenter transformed the wooden box into cup-task mode by covering the lever with the wooden occluder (see Figure 2). She pointed out to the participant that the box had changed and that the lever did not work any

https://doi.org/10.1017/S0954579401001079 Published online by Cambridge University Press

more. She placed the ball inside the box and lowered the screen. Next, she placed the cup on the table in front of the box then showed the participant, with accompanying explanation, that the only way to obtain the ball was to turn over the cup. She demonstrated how this method worked two times before the child was allowed to try, and also whenever there was an erroneous trial. Again, the criterion was three inversion-and-reach sequences in a row within 15 trials.

Finally, the participant was shown the aluminum box for the switch task. The experimenter placed the ball on the platform inside the box and then invited the participant to retrieve it. After he or she had failed to achieve this two times (because the direct reach triggered the trap-door mechanism), the experimenter drew the participant's attention to the switch. She explained, with a demonstration, that the only way to obtain the ball was to first throw the switch. The switch-and-reach sequence was demonstrated twice at the beginning and, as in the other tasks, each time the child failed a trial.

The procedures for the cup–switch–lever and switch–lever–cup conditions differed only in the following respects. In the cup–switch–lever order the wooden box was in the cup-task mode (the switch was already covered) at the beginning of the test session, so the participants did not know that there was another way of releasing the screen. It was expected that this would influence the type of errors participants would make in this task (see Scoring section below). A further difference was that in both the lever task and cup task participants were asked to retrieve the ball twice with the screen up at the beginning, which enabled them to observe that they could not get the ball out when the screen was lowered. Note again that the tasks were always explained both verbally and with demonstration in order to ensure that even children with low VMA fully understood the instructions.

Scoring

Each trial was coded as correct or as an error during testing. Only coordinated lever-switch-and-reach sequences in the lever task and switch task and inversion-and-reach sequences in the cup task were scored as correct. Additionally, we differentiated between five types of error and categorized them with the aid of the videotape records. A participant could be categorized as making only one type of error within a trial. The error types were as follows.

Direct-reaching (DR). The participant touched or pushed the screen (in the lever task or in the cup task) or made a direct reach that triggered the trap-door mechanism (in the switch task).

Nonreaching (NR). The participant did not reach after he or she had turned over the cup (in the cup task) or after he or she had thrown the switch or depressed the lever (in the other two tasks) within 5 s.

Other-route (OR). The participant tried to repeat a correct response from a previous task. For example, instead of turning over the cup he or she looked for the lever. This error could, in fact, occur only in the cup task under the lever–cup–switch or in the switch–lever–cup order.

Inappropriate-route (IR). The participant tried to manipulate the box in an inappropriate way that did not allow access to the ball. For example, he or she tried to rotate the screen back by the knob in the lever task and cup tasks or flicked the switch rapidly up and down in the switch task.

Pause (P). The participant did nothing for at least 5 s.

Results

All participants in the lever task attained the criterion. In addition, 96% of AUT and all MLD and normally developing children on the cup task did so. That is to say, they performed a sequence of three correct responses within 15 trials in these two tasks. However, only 70% of the AUT group reached criterion on the switch task as compared to 93% of the

https://doi.org/10.1017/S0954579401001079 Published online by Cambridge University Press

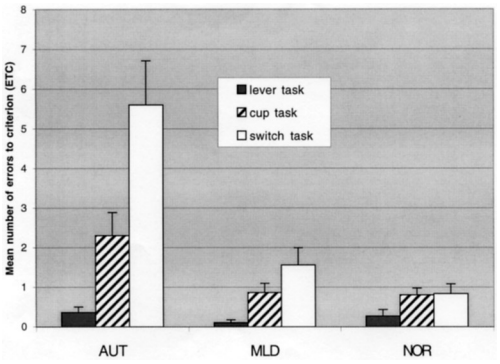


Figure 4. The figure shows the means and standard errors of the ETC (number of errors to criterion) in each task by each group ($N = 30$). AUT, children with autism; MLD, children with moderate learning difficulties; NOR, normally developing children.

MLD and 100% of the normally developing group. Comparison of the performance of the two clinical groups on the switch task a showed significant difference, $\chi^2 = 5.455$, $df = 1$, $p < 0.02$. Order of task presentation had no effect on performance on the switch task.

Errors to criterion

Performance was also analyzed in terms of the number of errors participants made before reaching the criterion or before the end of the session (errors to criterion [ETC]). As the performance of all groups was at ceiling on the lever task (see Figure 4), ETC scores of this task were not included in further analyses.

A three-way ANOVA was performed on ETC scores with the within-subject variable of task (2) and the between-subject variables of order (3) and group (3). The main effect of group was significant, F (2, 81) = 14,21 $p <$ 0.0001. There was also a significant main effect of task, F (2, 81) = 10.56, $p < 0.002$, and a significant interaction between task and group, F (2, 81) = 5,80, $p < 0.04$. The ETC scores are depicted in Figure 4. Because the ANOVA resulted in neither a main effect of order nor in any interaction involving the order factor, order was eliminated from further analyses.

Pairwise t test comparisons (with Bonferroni correction) were applied to the ETC data in order to compare performance on the two tasks within each group. In the AUT group a significant difference was found between the cup task and switch task, t (29) = –2.92, $p <$ 0.05. For the MLD and NOR groups, the ETC scores revealed no significant difference between the two tasks.

In order to compare the performance between groups on each task, separate one-way ANOVAs of the ETC scores in each group were also carried out, with the between-subjects variable of group (3). On the cup task, the ANOVA revealed a significant main effect of group, F (2, 87) = 4.95, $p < 0.01$. Post hoc analysis ($p < 0.05$ on Newman–Keuls) showed that the AUT group found the cup task significantly more difficult than did the other two groups. In the switch task, the main effect of group was clearly significant, F (2, 87) = 13.38, $p < 0.0001$. This was again due to the fact that the task was significantly more difficult for the AUT group than for the other two groups ($p < 0.05$ on Newman–Keuls).

These analyses can be summarized as follows. If children with autism found the switch task more difficult than the cup task, and their performance on both tasks was at a lower level than that of the other two groups.[2] For normally developing children and for children with moderate learning difficulties the switch task was not significantly harder than the cup task, and the performance of these two groups did not differ from each other in either of the two tasks.

Types of error

Table 1 shows the numbers of participants who committed different types of error in

2. Although the two clinical groups did not differ significantly from each other in terms of VIQ scores, the AUT group did have somewhat lower verbal IQs than the MLD group. Accordingly, separate ANOVAs were performed for each of the two tasks to covary VIQ. In the cup task, while this resulted in the main effect of group now failing to reach the conventional level of significance, there nonetheless remained a strong tendency towards significance, F (1, 57) = 2.85, $p <$ 0.063, indicating that the AUT group tended to have more difficulty with this task than did the other two groups. The main effect of group remained significant in the switch task, F (1, 57) = 4.66, $p < 0.035$.

https://doi.org/10.1017/S0954579401001079 Published online by Cambridge University Press

Table 1. *The number of participants committing different types of errors*

	Lever			Cup			Switch		
Error Types	AUT	MLD	NOR	AUT	MLD	NOR	AUT	MLD	NOR
Direct reach	4	1	0	4	1	0	19	12	11
Nonreaching	2	2	4	14	10	16	1	1	0
Other route	0	0	0	6	0	0	0	0	0
Inappropriate route	1	0	1	1	2	0	5	2	2
Pause	0	0	1	2	3	0	2	1	0

Note: Each cell represents the number of subjects who made a certain type of error in each group on each task. Note that a child could commit different types of errors in one particular task. AUT, children with autism; MLD, children with moderate learning difficulties; NOR, normally developing children.

each task. The largest proportion of participants who made errors made the DR error in the switch task. In this task, it can be seen that quite a number of children within each group made direct reaching errors. However, the AUT group differed from the other groups insofar as their members tended the make the error perseveratively. To illustrate, in the AUT group 14 participants made at least two consecutive direct reaches (6 of them made more than nine consecutive direct reaches), compared with 8 participants of MLD group and 5 of normally developing group.

While NR errors were almost as common as DR errors, they occurred typically in the cup task and were about equally distributed among the groups. Once more, children with autism tended to commit the error perseveratively: 43% of the AUT participants who made this error made at least two errors consecutively, compared with 31% of NOR group and only 10% of MLD group.

OR errors (e.g., attempting to operate the blocked lever) were found only in AUT group. While IR errors were not common in any of the groups, two participants in the AUT group produced a high number of them: 7 and 15 times in the switch task. P errors were found in only a handful children and were not specific to any group.

Finally, it is necessary to consider the commission of NR errors, for the following reason. It could be argued that NR errors have a special status insofar as they involve a partial success. Recall that an NR error is made when a participant correctly performs the means action (e.g., turning over the cup) but subsequently fails to perform the goal act (reaching for the ball). In contrast to this, the other four errors involve participants completely failing to perform the means act. In could be argued that performing the correct means act and then neglecting the goal is a failure of a different kind.

NR errors were most common in the cup task, as can be seen from Table 1. Indeed the only error that normally developing children made on the cup task was this error. It is unlikely that the commission of this kind of error was a symptom of participants' being unfamiliar with the testing situation because those children who made NR errors in the cup task were not more likely to have be given this task first ($\chi^2 = 1.16$, $df = 2$, $p > 0.05$).

In order to discover how committing the NR error affected group differences in ETC, a further analysis was run that excluded these errors. The modified ETC data were produced by simply subtracting the NR errors from the original ETC scores in each task. The new three-way ANOVA—Group (2) × Task (2) × Order–Condition (3)—revealed a main effect of group, F (2, 81) = 13.02; $p < 0.0001$, and a main effect of task, F (1, 81) = 25.04; $p < 0.001$. A significant interaction was found between the task and group factors, F (2, 81) = 7.06; $p < 0.001$, which is represented graphically in Figure 5.

Multiple t tests (with Bonferroni correction) on the relative task difficulty revealed that when NR responses were not treated as errors, the switch task continued to differ significantly in difficulty from the cup task within all groups: AUT, t (29) = −3.71, $p <$

https://doi.org/10.1017/S0954579401001079 Published online by Cambridge University Press

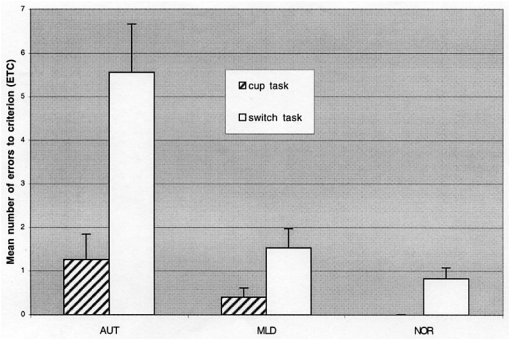


Figure 5. The figure shows the means and standard errors of the ETC (number of errors to criterion) in the cup task and the switch task by each group with the exclusion of NR (nonreaching) errors ($N = 30$). AUT, children with autism; MLD, children with moderate learning difficulties; NOR, normally developing children.

0.001; MLD, $t\ (29) = -3.22$, $p < 0.001$; NOR, $t\ (29) = -3.28$, $p < 0.001$. Post hoc, within-task comparisons between the performance of the groups showed that while the AUT group still performed at a lower level than the other two groups on the switch task ($p < 0.05$ Newman–Keuls) there was now no difference between the AUT group and the MLD group in the cup task on $p < 0.05$ level of significance on Newman–Keuls.

Discussion

As was to be expected, the children and adolescents with autism were not challenged by the lever task but were *considerably* challenged by the switch task. The latter finding replicates Hughes and Russell (1993, Experiment 2). Adequate performance on the lever task does, however, have to be treated with some caution, given that all groups were at ceiling on it. Indeed it might be argued that the fact that the box was a novel artifact rendered the means act novel and arbitrary—at least to some degree. What weakens this argument, however, is the fact that children from all three groups frequently invented the means before any instructions were given: they depressed the lever before the experimenter showed them how to retrieve the ball. Presumably, seeing the box activated a well-learned script such as "opening a door by a handle." Spontaneous invention of the means act never occurred in the cup task or in the switch task.

The adequate performance by children with autism in the lever task also excludes certain explanations for their relatively poor performance on the other tasks. First, since the linguistic complexity of the instructions was similar across the three tasks, the deficit cannot be attributed to problems with understanding the instructions. Second, because the means acts were demonstrated to the children by the experimenter, it might be argued that whatever difficulties were found among the children with autism are explicable in terms of their relatively poor mimetic capacities (Smith & Bryson, 1994). The fact that children with autism were not challenged by the lever task suggests they the difficulties that were in evidence elsewhere were not caused by lack of skill at imitation.

Broadly, the successful performance on the lever task by the children with autism can be said to be due to the fact that, despite its being a sensorimotor task, the lever task was not an executive task because there was no prepotency (a direct reach for the object was not possible) and the children did not have to update their model of the physical world or follow a rule.

Our main question was whether participants with autism would be impaired on the cup task, a task that did not encourage a prepotent response but that required the following an arbitrary means. With one qualification it can be said that the participants with autism did indeed perform at a level lower than both of the comparison groups on this task. We now turn to this qualification. The differences between the ETC scores of the participants with autism and those with MLD were reduced to below significance when the NR errors were excluded from the analysis. As we said earlier, it could be argued that these were not true errors, or at least not errors of the same status as the others. Moreover, what might have contributed to the fact that many participants failed to reach for the ball after inverting the cup was the fact that this somewhat magical event was distracting them from the goal. (There were often signs of surprise

https://doi.org/10.1017/S0954579401001079 Published online by Cambridge University Press

and pleasure.) If so—the skeptical view runs—it cannot be said that these were executive errors in the sense of errors caused by prepotency or by failure to follow a rule. Given this, if excluding errors that were not "true" executive removed group differences, then we cannot claim to have demonstrated group differences on this task that are informative about executive capacities in autism.

In reply, explaining the fact that many participants did not follow through after inverting the cup in terms of their becoming "distracted" is to do little more than describe the failure. It leaves unanswered the question of why the distraction was strong enough to lead to goal neglect. To illustrate, a parallel to the distraction argument would be dismissing Duncan's (1995) studies of goal neglect by saying that his participants were simply being "inattentive." Moreover, while it is reasonable to apply the "distractible" description to initial errors, it is a much less appropriate description of the way in which some participants (those with autism is particular) made the error repeatedly. Making the error more than once would seem to indicate a failure to appreciate that cup inversion was a means rather than an end in itself. Finally, in addition to, or instead of, taking delight in the procedure, some of the participants looked questioningly at the experimenter once they had raised the screen as if to say "Now what?" This suggests that some of the cup inversions had been simple acts of imitation rather than means acts. In sum, it is fair to assume that many participants who made NR errors did so because of a weak grasp of the fact that in this task an arbitrary means enables them to achieve a nonarbitrary goal.

With further regard to the question of whether NR errors were a form of executive error, they might indeed be regarded as themselves being indices of goal neglect. Perhaps there is a form of goal neglect prevalent among very young and mentally handicapped children in which the goallike properties of an action evaporate when the means act affording it becomes unintelligible. Be that as it may, we accept that our views on the status of NR error will seem contentious to some readers. We are currently seeking to clarify this question by running a version of the cup task with a more desirable reward (a sticker that the child can keep). In this case the number of NR errors might be reduced if their commission in this study was due to lack of motivation rather than to a lack of executive capacity.

We next consider why the participants with autism were much more strongly challenged by the switch task than by the cup task.[3] The obvious reason—this being part of the rationale for our study—would seem to be that the switch task afforded a prepotent response and the cup task did not. But it is possible that the switch task's containing a prepotent lure was not the reason for the difference, or at least that it was not the only reason for it. To explain, in contrast to the means act in the switch task, the means act in the cup task *had an immediate and perceptible effect* on the physical world: the inversion of the cup made the screen pop up. In the switch task, on the other hand, throwing the switch appeared to make nothing happen. One might say indeed that this fact *reduced the arbitrariness* of the means action in the cup task. For unlike following the rules of a game, following the cup-inversion rule made "goods and services" immediately available. Consistent with the claim that inhibition problems may not have been responsible for the differences in switch–cup task difficulty is the fact that research by Ozonoff and her coworkers has shown that on the Go–NoGo task (Ozonoff, Strayer, McMahon, & Filloux, 1994) and on the Stopping and the Negative Priming tasks (Ozonoff & Strayer, 1997)—all three tests of inhibition—children with autism are not generally impaired. It is also relevant to note that the children with autism in Hughes and Russell's (1993) study performed at a similar level to controls in a variant of their task in which the

3. It is necessary to point out that the lack of a significant difference in performance on these two tasks in the case of the MLD children was almost certainly due to the fact that performance on both tasks was so good in this group that differences could not be revealed. In any event, our assumptions are that (a) *any* population of children will be challenged more by tasks that include prepotency (as did the switch task) and (b) that this will be particularly true of children with autism.

https://doi.org/10.1017/S0954579401001079 Published online by Cambridge University Press

object had to be retrieved by turning a knob on the side of the box. In this case there was a direct, causal link between a turn of the knob and a paddle, which was visible inside the box, knocking the object (a marble) down a chute and out. This tells us that the children with autism were able to overcome the effects of prepotency so long as the procedure they had to follow presented a visible and intelligible link between act and outcome.

The fact that Hughes and Russell's task was much easier when the act–outcome link was intelligible and visible raises the possibility that when children succeeded on the switch version they did so because they could draw inferences about unseen mechanical processes. That is to say, they were able to construct a representation of the invisible result of the means action and infer nonobservable changes in the state of the box. Indeed it is not inconceivable that this ability to draw inferences about unperceived processes may be a general problem in autism, given that so-called "mentalizing" tasks like the false belief task also require inferences to be drawn about processes going on, unperceived, "in" the mind. Autistic children's apparently intact abilities in "folk physics" might seem to contract this view, but it should be noted that the tasks used to assess these do not require inferences to be made about unobservable mechanical processes. For example, children may be required to sequence pictures in which a running girl is caused to fall over by a brick in her path (see Baron–Cohen, 1997, for a review of the evidence for intact or superior folk physics in autism). Moreover, the claim that children with autism are not challenged by the false photograph task (a mechanical version of the false belief task) is open to question (Russell, Saltmarsh, & Hill, 1999).

Whatever the cogency of ideas of this kind, our study has at least yielded suggestive evidence of an autism-specific difficulty with means–end tasks in which there is no prepotent lure but in which the means act is arbitrary in nature. Before we consider the implications of this result, we must mention two respects in which it needs to be qualified. In the first place, both of the clinical groups had moderate levels of retardation and so it is possible that this difficulty is specific to low-functioning children with autism an so may not generalize to children with normal intelligence. Second, note that the difference between the performance of autistic and the MLD group on the cup task was eroded to a level just below significance when IQ was covaried from the analysis ($p < 0.063$; see footnote 2). The fact that both removal of NR errors and the covarying of verbal IQ affected group differences ensures that no secure conclusions can be drawn on this matter in the absence of further research.

We finally consider how these difficulties with arbitrary rule following (emerging in this study and perhaps in future research) can be located in relation to more mainstream ideas about the nature of executive dysfunctions in autism. One such "mainstream" autistic executive impairment is with set shifting (see Ozonoff, 1997, for a review).

First, it is important to note that children with autism are perfectly able to learn to follow novel and arbitrary *instructions*—the direct instruction to "Do *X*" when no other tendency is competing. For instance, in the first phase of the WCST they have no difficulty in sorting the cards by color, shape, or number. Their difficulty arises when they have to follow a different rule. As we have been arguing, this naturally suggests that children with autism only experience difficulty when there is both a prepotent lure (the well-entrenched "sort-by-*X*" rule) and arbitrariness. But we have shown that even when no prepotent response is possible, executing an arbitrary means presents difficulties. We can bridge these two areas of work (on set shifting and on taking arbitrary actions) with the following account. In the cup task, children are surely bringing to the experiment certain presuppositions about what one can do and what is meant to be done with cups. That is to say, the cup presents the child with a certain, "functional frame," as we will call it, something that can be understood in terms of Gibson's notion of "affordances" (Gibson, 1979). The functional frame of a cup subsumes affordances such as drinking from, pouring from, covering with, and so forth. There is no linkage, however, between these set of afford-

https://doi.org/10.1017/S0954579401001079 Published online by Cambridge University Press

ances and those of a now open box containing a desired object, which affords reaching into, retrieving *X*, and playing with *X*. Such affordances make up the functional frame of the box and its contents. In the light of this, we can regard the challenge to the subject as being that of moving between sets of affordances, and so *in this sense the cup task is a set-shifting task embedded in a means–end context*. The notion of a functional frame is the executive equivalent to a "set" in a set-shifting task.

This idea can be tested in at least the following two ways. In the first place, if the object that has to be manipulated is not something with a familiar function (unlike a cup) the group differences should vanish. Second, and more generally, the idea that children with autism find it hard to move between functional frames makes the prediction that they will be more challenged than are control children by manipulations of functional fixedness (Duncker, 1945). Functional fixedness experiments show that subjects have difficulty appreciating that an object can be used as, say, a weight if they had previously been using it as a transformer: they are functionally fixated on it under the prior aspect. Successful action requires them to move from the functional frame of electrical equipment to the functional frame of weighing objects down for a specific purpose. The prediction is that children with autism will find it especially difficult to move between functional frames. We are currently investigating this possibility.

References

Abrahamsen, E. P., & Mitchell, J. R. (1990). Communication and sensorimotor functioning in children with autism. *Journal of Autism and Developmental Disorders, 20*, 75–85.

Adrien, J. J., Marineau, J., Barthélémy, C., Bruneau, N., Garreau, B., & Savage, G. (1995). Disorders of regulation of behavior and cognitive activity in autistic children. *Journal of Autism and Developmental Disorders, 25*, 249–263.

Ahmed, A., & Ruffman, T. (1998). Why do infants make A not B errors in a search task, yet show memory for the location of hidden objects in a nonsearch task? *Developmental Psychology, 34*, 441–453.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders (3rd ed., Rev.)*. Washington, DC: Author.

Baillargeon, R. (1993). The object concept revisited: New directions in the investigation of infants' physical knowledge. In C. E. Granrud (Ed.), *Visual perception and cognition in infancy: Carnegie–Mellon Symposia on Cognition* (pp. 265–325). Hillsdale, NJ: Erlbaum.

Baron–Cohen, S. (1997). Are children with autism superior at folk physics? In H. Wellman & K. Inagaki (Eds.), *The emergence of core domains of thought: Children's reasoning about physical, psychological, and biological phenomena. New Directions for Child Development,* San Francisco: Jossey–Bass.

Baron–Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a theory of mind. *Cognition, 21*, 37–46.

Diamond, A. (1985). Development of the ability to use recall to guide action, as indicated by infants' performance on A-not-B. *Child Development, 56*, 868–883.

Diamond, A. (1988). Abilities and neural mechanisms underlying A-not-B performance. *Child Development, 59*, 523–527.

Duncan, J. (1995). Attention, intelligence, and the frontal lobes. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences*. Cambridge, MA: MIT Press.

Duncker, K. (1945). On problem solving. *Psychological Monographs, 58*, 5.

Dunn, L. M., Dunn, L. M., Whetton, C., & Pintilie, D. (1982). *British Picture Vocabulary Scale*. Windsor, UK: NFER–Nelson.

Frith, U., & Happé, F. (1994). Autism: Beyond "theory of mind." *Cognition,* 50, 115–132.

Gibson, J. (1979). *The ecological approach to visual perception.* Boston: Houghton Mifflin Company.

Hood, B. (1995). Gravity rules for 2–4 year olds. *Cognitive Development, 10*, 375–386.

Hughes, C., and Russell, J. (1993). Autistic children's difficulty with mental disengagement from an object: Implications for theories of autism. *Developmental Psychology, 29*, 498–510.

Hughes, C., Russell, J., & Robbins, T. W. (1994). Evidence for executive dysfunction in autism. *Neuropsychologia. 32*, 477–492.

McEvoy, R. E., Rogers, S. J., & Pennington, B. F. (1993). Executive function and social communication deficits in young, autistic children. *Journal of Child Psychology and Psychiatry, 34*, 563–578.

Ozonoff, S. (1997). Components of executive function in autism and other disorders. In J. Russell (Ed.), *Executive functioning in autism*. Oxford: Oxford University Press.

Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1991). Executive deficits in high-functioning autistic children: Relationship to theory of mind. *Journal of Child Psychology and Psychiatry, 32*, 1081–1105.

Ozonoff, S. & Strayer, D. L. (1997). Inhibitory function in nonretarded children with autism. *Journal of Autism and Developmental Disorders, 27*, 59–77.

Ozonoff, S., Strayer, D. L., McMahon, W. M., & Filloux, F. (1994). Executive function abilities in autism and Tourette syndrome: An information processing approach. *Journal of Child Psychology and Psychiatry, 35*, 1015–1032.

Pennington, B., Rogers, S., Bennetto, L., Griffith E. M., Reed, D. T., & Shyu, V. (1997). Validity tests of the

https://doi.org/10.1017/S0954579401001079 Published online by Cambridge University Press

executive dysfunction hypothesis of autism. In J. Russell (Ed.), *Executive functioning in autism.* Oxford: Oxford University Press.

Piaget, J. (1954). *The child's construction of reality*. London: Routledge and Kegan Paul.

Rumsey, J. M. (1985). Conceptual problem-solving in highly verbal, nonretarded autistic men. *Journal of Consulting and Clinical Psychology, 48*, 605–614.

Rumsey, J. M., & Hamburger S. D. (1988). Neuropsychological findings in high-functioning with infantile, autism, residual state. *Journal of Clinical and Experimental Neuropsychology, 10*, 201–221.

Russell, J. (1997a). *Autism as an executive disorder*. Oxford: Oxford University Press.

Russell, J. (1997b). How executive disorders can bring about an inadequate "theory of mind." In J. Russell (Ed.), *Autism as an executive disorder*. Oxford: Oxford University Press.

Russell, J., Jarrold, C., & Hood, B. (1999). Two intact executive capacities in children with autism: Implications for the core executive dysfunctions in the disorder. *Journal of Autism and Developmental Disorders, 29*, 103–112.

Russell, J., Saltmarsh, R., & Hill, E. (1999).What do executive factors contribute to the failure on false belief tasks by children with autism? *Journal of Child Psychology and Psychiatry, 40*, 859–867.

Smith, I. M., & Bryson, S. E. (1994). Imitation and action in autism: A critical review. *Psychological Bulletin, 116*, 259–273.

Wehner, E., & Rogers, S. (1994, March). Attachment relationship of autistic and developmentally delayed children. Paper presented at the bimonthly meeting of the Developmental Psychobiology Research Group, Denver, CO.

https://doi.org/10.1017/S0954579401001079 Published online by Cambridge University Press