Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Journal of Experimental Child Psychology 98 (2007) 184–192
Journal of Experimental Child Psychology
www.elsevier.com/locate/jecp

Brief Report

Choosing between two objects reduces 3-year-olds' errors on a reverse-contingency test of executive function

Daniel J. Carroll [a,*], Ian A. Apperly [b], Kevin J. Riggs [c]

[a] *Department of Psychology, University of Sheffield, Western Bank, Sheffield S10 2TN, UK*
[b] *School of Psychology, University of Birmingham, Edgbaston B15 2TT, UK*
[c] *Department of Psychology, London Metropolitan University, Old Castle Street, London E1 7NT, UK*

Received 13 May 2007; revised 28 July 2007
Available online 18 September 2007

Abstract

In the present experiment, we used a reversed-contingency paradigm (the windows task: [Russell, J., Mauthner, N., Sharpe, S., & Tidswell, T. (1991). The windows task as a measure of strategic deception in preschoolers and autistic subjects. *British Journal of Developmental Psychology, 9*, 331–349]) to explore the effect of alterations in the task array on 3-year-old children's strategic reasoning. Children were offered a choice between either a desirable object and an undesirable object, or between a desirable object and an empty location. There was significantly better performance on the two-object version of the task. This difference was evident even on subsequent trials when the second object was removed and the empty location reintroduced. This suggests that presenting children with a choice between two objects helps them to formulate a strategy, rather than to execute a previously determined response.

Keywords: Executive function; Strategic reasoning; Inhibition; Theory of mind; Metacognition

During the preschool years, children's cognitive abilities undergo substantial changes across a range of domains. Many of these changes have been explained in terms of the emergence of improved executive functions, the basic cognitive processes that, singly or

* Corresponding author. Fax: +44 114 276 6515.
E-mail address: d.carroll@sheffield.ac.uk (D.J. Carroll).

0022-0965/$ - see front matter
doi:10.1016/j.jecp.2007.08.001

in concert, allow agents to produce flexible, goal-oriented behavior (Klenberg, Korkman, & Lahti-Nuutila, 2001; Welsh, Pennington, & Groisser, 1991). A variety of tasks have been developed to look at individual executive functions; one widely used paradigm that directly assesses children's ability to produce flexible, goal-oriented behavior is the reversed-contingency task. There have been many versions of this type of task, conducted both with children (e.g., the windows task, Russell, Mauthner, Sharpe, & Tidswell, 1991; the Less Is More task, Carlson, Davis, & Leach, 2005; the deceptive box task; Carlson, Moses, & Hix, 1998) and with primates (e.g., Boysen & Berntson, 1995; Kralik, Hauser, & Zimlicki, 2002; Vlamings, Uher, & Call, 2006). These tasks vary in their specific demands and instructions, but in all cases participants try to obtain a desirable reward by pointing to an array with two locations; the option they point to is taken away, and participants receive the option they did not point to. Successful performance therefore requires that the participant is able to formulate and act upon the basic strategy of pointing away from a desirable object in order to receive it.

A reliable developmental finding is that 3-year-old children find it difficult to win the desirable object by pointing away from it. There has, nevertheless, been marked variability in children's performance. By far the poorest performance was found in Russell et al.'s (1991) study, in which 16 out of 17 children failed to point to the less desirable option on the first trial. Strikingly, 11 of these children went on to fail on all 20 of their experimental trials. This severe degree of difficulty has been replicated in a subsequent study (Hala & Russell, 2001) but not in all studies (e.g., Carlson et al., 2005; Carroll, Apperly, & Riggs, 2007; Russell, Hala, & Hill, 2003; Samuels, Brooks, & Frye, 1996; Simpson, Riggs, & Simon, 2004). Better performance in some of these studies can be attributed to children using easier, non-standard response modes (e.g., Carlson et al., 1998; Russell et al., 2003) or the use of wording that makes it easier for children to work out how to respond (Carroll et al., 2007; Samuels et al., 1996; Simpson et al., 2004). Nevertheless, these explanations cannot account for all findings. In Carlson et al.'s (2005) reversed-contingency task – as in Russell et al. (1991) – children had to work out the appropriate response and then indicate their chosen option by pointing with their hands. In the Carlson study, however, even young 3-year-olds succeed on almost half of their trials (49%), and older 3-year-olds performed even better (61%). So whilst succeeding with a reversed contingency was not trivially easy for these children, they performed remarkably well.

One potentially crucial difference between these studies is the array that children have to act upon. The Less Is More task (Carlson et al., 2005) presents children with a choice between different quantities of desirable objects. In contrast, the windows task (Russell et al., 1991) involves a choice between a desirable object at one location, and nothing (a visibly empty box) at the other location. This presence or absence of an empty location may be key. Placing a marker at an otherwise empty location may help children to execute a response by providing a target for their action (see e.g., Bekkering, Wohlschlager, & Gattis, 2000, and Brodeur, 2004). Alternatively, having a less desirable alternative to the object that the child wishes to win may encourage children to think about the task in a different way. It is already known that children perform well on reversed-contingency tasks when they are told to "point to the empty box" (Simpson et al., 2004) or to "point to the box for [your opponent] to open so that [he] doesn't get the sticker" (e.g., Carroll et al., 2007). The presence of a less desirable object may help children to succeed in a similar way to these wording manipulations by enabling children to think about which box to give away rather than which box they wish to obtain. These explanations make different predictions about the nature of any benefit children might gain from the addition of a less desirable object to the standard windows task, and it is these predictions that are tested in the current experiment.

We compared children's performance on a standard windows task array (a choice between a reward and an empty location) with performance on an array where the incorrect response was represented by an undesirable object (a choice between a reward and a small scrap of paper) to see whether eliminating the empty location improves children's strategic reasoning. The task logic in both conditions was the same – children must point to the non-desired option to receive the reward item – and any performance differences would hence be attributable to the difference in the array.

We also asked a second question: if a second object does help children, *how* does it help them? Does it enable children to work out how they should respond or to overcome inhibitory demands on the task? To address this question, we included three experimental conditions: we compared a Single Object condition (in which one box contains a reward, and the other box is empty – the standard windows task array); a Two Object condition (one box contains a reward, the other contains a scrap of paper); and a Mixed condition (identical to the Two Object condition for the first three trials, and then identical to the Single Object for the remaining trials). This Mixed condition allowed us to see whether any benefit from the Two Object array sustained after the second object is withdrawn. This kind of transfer manipulation has previously been used with both children and primates, with conflicting findings. Boysen and Berntson (1995) found that chimpanzees' good performance in a symbol condition deteriorated when the symbols were withdrawn. In contrast, Apperly and Carroll (2006)[1] found that children's good performance in symbol conditions was sustained even after symbols were removed, and Carroll et al. (2007) found similar sustained improvement with alternative task wordings that facilitated children's performance.

In the current study we expected one of three possible outcomes. The first possibility was that children did not find it any easier to respond to a two object array than to a single object array, in which case performance in all three conditions would be poor. A second possibility was that the second object helped children with the trial-by-trial demands of executing a response. If this were the case, children would perform well in the Two Object condition, and equally well for the first three trials of the Mixed condition; their performance would then drop back to Single-Object levels when the two-object array was replaced with a single-object array. This was the pattern of responding that Boysen and Berntson (1995) found when chimpanzees were presented with a reverse-contingency task that used symbols. The third possibility was that the second object helped children to work out how they should respond. In this case, performance in the Two Object condition would be good, and performance in the Mixed condition would be good and would remain better than the Single Object condition *even when* the single-object array was introduced. In this case we would conclude that the initial exposure to the second object had given children a basis for responding that allowed them to overcome any difficulties caused by the single-object array.

[1] A written report of these findings is under submission and available from the authors on request.

Method

Participants

Sixty-eight children attending nursery schools in lower-middle class areas of the UK participated (39 boys). The mean age of the sample was 3:7 years (range 3:1 to 4:0). Children were randomly assigned to one of three experimental conditions. A one-way analysis of variance confirmed that the mean ages of the groups were not significantly different ($F(2,65) = 0.56$, $p = 0.57$).

Materials

Two boxes (one red, one blue) measuring 9 cm × 9 cm × 12 cm were used for the training phase. For the testing phase, two similar boxes with windows cut in the side of them were used. For the second object in the Two Object condition, scraps of paper were used.

Procedure

Participants were tested individually by a male experimenter in a room adjacent to the main classroom. The child and the experimenter sat facing each other across a table. Children underwent a training phase to ensure that they understood the task rules and the method of responding. There were three experimental conditions: a Single Object condition, a Two Object condition, and a Mixed condition.

Training phase – Single Object condition

Children were shown the non-windowed boxes, and were told that the experimenter was going to put a treat in one of the boxes. They were told that they would choose which box the experimenter looked in. It was explained that if the experimenter found the treat he would keep it, but that if the experimenter didn't find it, the child would get to open the other box and keep the treat. The treat was hidden out of view of the child, and children were asked to point to a box for the experimenter to open. If the experimenter found the treat, he kept it; if he failed to find it, the second box was opened and the child got to keep it. This was repeated for five training trials. On the sixth trial, after the indicated box had been opened, the child was asked "Who gets the treat this time?" to ensure that the child understood the task rules. This questioning was repeated on subsequent trials until the child had given three consecutive correct responses. All children managed this within 10 trials.

Training phase – Two Object and mixed conditions

The procedure was similar to the Single Object condition, except that on each training trial, a treat was hidden in one box, and a small scrap of paper was hidden in the other box. On every trial, the experimenter opened the box that the child pointed to, and then the child opened the remaining box. The scrap of paper was always discarded after each trial, regardless of whose box it was in.

Testing phase – Single Object condition

The training boxes were replaced with the windowed boxes. The windows were pointed out, and the experimenter observed that that the child would now be able to see where the treat was. A treat was placed in one of the boxes and the child was prompted to indicate a box for the experimenter with the words "Point to a box for me [i.e., the experimenter] to open." This continued for 9 trials. After each trial the experimenter announced "I keep the treat this time," or "You keep the treat this time," as appropriate.

Testing phase – Two Object condition

The procedure was similar to the Single Object condition, but on each trial one box was baited with a treat, and the other box was baited with a scrap of paper. On each trial, the experimenter opened the box indicated by the child and removed the contents, after which the child opened the other box and removed its contents. Whoever found the treat kept it; whoever found the piece of paper discarded it.

Testing phase – Mixed condition

For the Mixed condition testing phase, trials 1 to 3 were identical to the Two Object condition testing phase. After trial 3, the experimenter said "We don't need to use the bits of paper anymore." Trials 4 to 9 were identical to the Single Object condition testing phase.

Results

Mean performance on each test trial of each condition is plotted in Fig. 1. To study the effects of learning from feedback, individual scores from each condition were

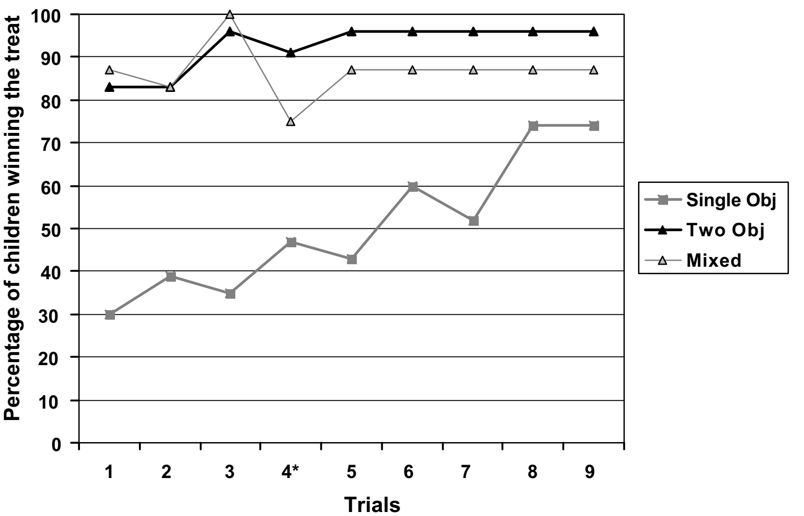

Fig. 1. Percentage of children responding correctly on each trial of the windows task for each condition. (* denotes the switch trial in the Mixed condition).

grouped into three-trial epochs and entered into an ANOVA, with condition as a between-subject factor, and epoch as a within-subject factor. This revealed a main effect of condition, $F(2,195) = 33.68$, $p < 0.001$, $\eta^2 = 0.27$, and a near-significant effect of epoch, $F(2,195) = 2.69$, $p = 0.071$, $\eta^2 = 0.03$. Post-hoc Bonferroni comparisons found performance in the Single Object condition to be significantly worse than both the Two Object and Mixed conditions, both $p < 0.001$, but found no significant difference between the Two Object and Mixed conditions, $p = 0.62$. Bonferroni post-hoc tests found that children's performance on the third epoch tended towards being better than performance on the first epoch, $p = 0.061$; performance on the second epoch did not significantly differ from either the first, $p = 0.99$, or third epoch, $p = 0.47$. The interaction between epoch and condition tended towards significance, $F(4,195) = 1.98$, $p = 0.10$, $\eta^2 = 0.04$, which may reflect the gradual improvement over time in the Single Object condition contrasted with the near-ceiling performance in the Two Object and Mixed conditions (see Fig. 1).

A chi-square comparison found that first trial performance differed across conditions, $x^2(2) = 19.91$, $p < 0.001$ (see Table 1). Follow-up comparisons (with Bonferroni-corrected alpha values of 0.016 denoting significance) found that first trial performance in the Single Object condition was worse than both the Two Object condition, $x^2(1) = 12.74$, $p < 0.001$, and the Mixed condition, $x^2(1) = 14.42$, $p < 0.001$, suggesting that the benefit conveyed by the second object arises before children have received any feedback. There was no difference between the Two Object and Mixed conditions, $x^2(1) = 0.12$, $p = 0.73$.

To look at performance on the switch trial in the Mixed condition, a chi-square analysis of trial 4 performance was conducted and found significant differences between conditions, $x^2(2) = 10.54$, $p = 0.005$. Follow-up comparisons (with Bonferroni-corrected alpha values of 0.016 denoting significance) found that trial 4 performance in the Single Object condition was worse than in the Two Object condition, $x^2(1) = 10.27$, $p = 0.001$. Performance in the Mixed condition did not significantly differ from either the Single Object, $x^2(1) = 2.91$, $p = 0.09$, or Two Object, $x^2(1) = 2.66$, $p = 0.10$, conditions. A Fisher's Exact Test to investigate the change in performance within the Mixed condition between trial 3 and trial 4 found that the decline in performance was significant, $p = 0.004$.

Mean performance across trials 4 to 9 was compared to see whether children in the Mixed condition continued to perform well after the second object was removed. A one-way ANOVA found significant differences between conditions, $F(2,65) = 8.82$, $p < 0.001$. Bonferroni post-hoc comparisons found that performance in the Single Object condition was significantly worse than both the Two Object condition, $p < 0.001$, and the Mixed condition, $p = 0.022$. There was no significant difference between the Two Object and Mixed conditions, $p = 0.59$.

Table 1
Mean performance on the windows task for each condition

	N	Overall/9	1st trial	Trials 1–3	Trial 4	Trials 4–9
Single Object	23	4.6 (51%)	7/23 (30%)	1.0 (35%)	11/23 (48%)	3.5 (59%)
Two Object	23	8.3 (92%)	19/23 (83%)	2.6 (87%)	21/23 (91%)	5.7 (95%)
Mixed	22	7.7 (86%)	19/22 (87%)	2.7 (90%)	16/22 (73%)	5.0 (84%)

Discussion

As expected, performance in the Single Object condition – the standard array for the windows task – was poor. In contrast, performance in the Two Object and Mixed conditions was significantly better. Moreover, in both the Two Object and Mixed conditions, children performed well from the very first trial, suggesting that the difference between conditions cannot be a consequence of children in these conditions learning more effectively from feedback.

A comparison of performance on trials 4 to 9 showed that children in the Mixed condition performed significantly better than children in the Single Object condition, even though the experimental procedures were identical for these trials. In the Mixed condition, the initial exposure to the second object appears to lead to better performance on subsequent trials where the single-object array was used. Withdrawing the manipulation in the Mixed condition did not, therefore, lead children to revert back to the poor performance of the Single Object condition. Some children in the Mixed condition showed a dip in performance on trial 4 (their performance on this trial was significantly worse than the previous trial, but did not differ from either standard or Two-Object condition performance). However, the rapid recovery in performance on subsequent trials suggests that this single-trial poor performance is likely to be the result of a switch-cost, where children have to take account of the new task conditions.

These results strongly support the notion that children do much better on reversed-contingency tasks that present them with a choice between two items, rather than a single item and an empty location. This suggests that much of the difference in performance between the Russell and Carlson versions of the task may be attributed to the different arrays used. Of course, this observation does not in any way diminish the findings drawn from these paradigms. Rather, it highlights a new and previously neglected task factor, and emphasizes that the type of choice children face must be taken into account when evaluating their performance. Interestingly, the same may not be the case for chimpanzees, who have tended to perform poorly despite all existing studies using reverse-contingency tasks in which the animals made a choice between two items or quantities of differing value (e.g., Boysen & Berntson, 1995; Kralik et al., 2002; Vlamings et al., 2006).

The current study also tested whether any benefit from the array was helping children to execute a response, or to think about the task differently. We found that the second object acts to enable children to work out a basis for responding, and not to free them from an executive error of pointing at a desirable object. This is importantly different to findings with reversed-contingency tasks used with primates: Boysen and Berntson (1995) found that although chimpanzees' poor performance could be boosted (through symbolic representation of the reward items), performance declined when the original task array was reintroduced, suggesting that chimpanzees' problems were with executing a response on the standard array. The present study shows a qualitatively different pattern in children than in chimpanzees, suggesting that ostensibly similar tasks pose very different challenges.

The present study suggests that children's ability to act flexibly is constrained not merely by the logic of the task (see e.g., Zelazo & Frye, 1998), nor by the means with which they respond (e.g., Carlson et al., 1998), nor by the "strength" of their representation of the task (e.g., Munakata, 2001), but also by how the specific instantiation of the task encourages children to approach the problem. We suggest that a Single Object array leads children to conflate the object they want (the baited box) and the object they must act upon, and their error is to use their overall goal (gaining the reward) as a basis for their immediate action. We suggest that the Two Object array encourages children to formulate a new strategy of selecting the item to give away. If children's goal is to select the item to give away, then there is unlikely to be any prepotent tendency to point at (and hence lose) the reward. This new strategy would work equally well to bypass the executive demands of the standard Single Object task, so explaining children's ability to transfer their improved performance from the Two Object task to the Single Object task.

Interestingly, the same transfer of improved performance has been observed in a Single Object version of the task when children are assisted with a change of wording (Carroll et al., 2007) or by replacing objects with symbols (Apperly & Carroll, 2006), suggesting that children may have been prompted to formulate a new response strategy in these cases too. This growing body of evidence raises an interesting question about how older children *spontaneously* succeed on standard reverse-contingency tasks. One possibility is that older children simply have better inhibitory control than younger children. Another option, highlighted by the current findings, is that older children *spontaneously* formulate a response strategy that bypasses key demands of the task.

References

Apperly, I. A., & Carroll, D.J . (2006). How do symbols affect children's strategic reasoning on a reverse-contingency task? Paper presented to the Experimental Psychology Society, 10–12 July, Plymouth, UK.

Bekkering, H., Wohlschlager, A., & Gattis, M. (2000). Imitation of gestures in children is goal-directed. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 53*(1), 153–164.

Boysen, S. T., & Berntson, G. G. (1995). Responses to quantity – perceptual versus cognitive mechanisms in chimpanzees (Pan troglodytes). *Journal of Experimental Psychology – Animal Behavior Processes, 21*(1), 82–86.

Brodeur, D. A. (2004). Age changes in attention control: assessing the role of stimulus contingencies. *Cognitive Development, 19*(2), 241–252.

Carlson, S. M., Davis, A. C., & Leach, J. G. (2005). Less is more – executive function and symbolic representation in preschool children. *Psychological Science, 16*(8), 609–616.

Carlson, S. M., Moses, L. J., & Hix, H. R. (1998). The role of inhibitory processes in young children's difficulties with deception and false belief. *Child Development, 69*(3), 672–691.

Carroll, D. J., Apperly, I. A., & Riggs, K. J. (2007). The executive demands of strategic reasoning are modified by the way in which children are prompted to think about the task: Evidence from 3- to 4-year-olds. *Cognitive Development, 22*(1), 142–148.

Hala, S., & Russell, J. (2001). Executive control within strategic deception: A window on early cognitive development? *Journal of Experimental Child Psychology, 80*(2), 112–141.

Klenberg, L., Korkman, M., & Lahti-Nuutila, P. (2001). Differential development of attention and executive functions in 3- to 12-year-old Finnish children. *Developmental Neuropsychology, 20*(1), 407–428.

Kralik, J. D., Hauser, M. D., & Zimlicki, R. (2002). The relationship between problem solving and inhibitory control: Cotton-top tamarin (*Saguinus oedipus*) performance on a reversed-contingency task. *Journal of Comparative Psychology, 116*(1), 39–50.

Munakata, Y. (2001). Graded representations in behavioral dissociations. *Trends In Cognitive Sciences, 5*(7), 309–315.

Russell, J., Hala, S., & Hill, E. (2003). The automated windows task: the performance of preschool children, children with autism, and children with moderate learning difficulties. *Cognitive Development, 18*(1), 111–137.

Russell, J., Mauthner, N., Sharpe, S., & Tidswell, T. (1991). The windows task as a measure of strategic deception in preschoolers and autistic subjects. *British Journal of Developmental Psychology, 9*, 331–349.

Samuels, M. C., Brooks, P. J., & Frye, D. (1996). Strategic game playing in children through the windows task. *British Journal of Developmental Psychology, 14*, 159–172.

Simpson, A., Riggs, K. J., & Simon, M. (2004). What makes the windows task difficult for young children: Rule inference or rule use? *Journal of Experimental Child Psychology, 87*(2), 155–170.

Vlamings, P., Uher, J., & Call, J. (2006). How the great apes (Pan troglodytes, Pongo pygmaeus, Pan paniscus, and Gorilla gorilla) perform on the reversed-contingency task: The effects of food quantity and food visibility. *Journal of Experimental Psychology – Animal Behavior Processes, 32*(1), 60–70.

Welsh, M. C., Pennington, B. F., & Groisser, D. B. (1991). A normative developmental study of executive function: a window on prefrontal function in children. *Developmental Neuropsychology, 7*(2), 131–149.

Zelazo, P. D., & Frye, D. (1998). Cognitive complexity and control: the development of executive function. *Current Directions in Psychological Science, 7*, 121–126.