Journal of Experimental Child Psychology 112 (2012) 312–325

Contents lists available at SciVerse ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

How do alternative ways of responding influence 3- and 4-year-olds' performance on tests of executive function and theory of mind?

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

[a] *Department of Psychology, University of Sheffield, Western Bank, Sheffield S10 2TN, UK*
[b] *Department of Psychology, University of Hull, Hull HU6 7RX, UK*
[c] *School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK*

ARTICLE INFO

Article history:
Received 26 May 2011
Revised 2 March 2012
Available online 6 April 2012

Keywords:
Executive function
Pointing
Inhibition
Theory of mind
Reasoning
Response mode

ABSTRACT

A total of 69 preschool children were tested on measures of false belief understanding (the Unexpected Transfer task), inhibitory control (the Grass/Snow task), and strategic reasoning (the Windows task). For each task, children indicated their response either by pointing with their index finger or by using a nonstandard response mode (pointing with a rotating arrow). The means of responding had no effect on children's performance on the Grass/Snow task or on the Unexpected Transfer task, although children performed better on the Unexpected Transfer task when the key object in the story was removed. In contrast, performance on the Windows task was significantly better when children pointed with the rotating arrow. A follow-up experiment with 79 preschoolers found that this improved performance on the Windows task was sustained even after the nonstandard response mode was removed and children again pointed with their finger. These findings together suggest that nonstandard response modes do not help children to inhibit prepotent pointing responses but may help them to formulate response strategies on reasoning tasks by discouraging unreflective impulsive responding.

---

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

http://dx.doi.org/10.1016/j.jecp.2012.03.001

Introduction

It is widely recognized that 3-year-olds have great difficulty in situations that require them to make a response that goes against their initial tendency or inclination. For example, children of this age have difficulty in postponing immediate gratification in order to obtain a larger reward later (e.g., Mischel, Shoda, & Rodriguez, 1989; Moore, Barresi, & Thompson, 1998). They have difficulty in controlling their tendency to give away the right answer in a game where their task is to deceive an opponent (e.g., Sodian, 1994). They find it difficult to respond to questions about other people's knowledge states when those states do not match up with their own (e.g., Wellman, Cross, & Watson, 2001). They also have substantial difficulty in learning to point to an empty box in order to be left with a second box containing a reward even over 20 repeated trials (Russell, Jarrold, & Potel, 1994). Understanding how children learn to overcome these problems to produce flexible goal-directed behavior informs us about their developing executive function. One way that has been found to help children overcome these difficulties is by changing the way in which they indicate their response. In the current study, we investigated the means by which nonstandard response modes (e.g., indicating a response by rotating an arrow rather than pointing with the hand) enhance children's ability to act on tasks that require them to respond in a counterintuitive manner.

One particularly informative paradigm in this regard is the Windows task, a test of children's strategic reasoning developed by Russell, Mauthner, Sharpe, and Tidswell (1991). Preschool children were presented with two closed boxes, one of which contained a treat (a candy or a sticker). During a training phase, children were taught that whichever box they pointed to would be taken away and given to an opponent, leaving them to open the box that remained. Once children had shown that they understood this rule, the opaque boxes were replaced with boxes that had windows cut into the side. By looking through the windows, children were able to see which box contained the treat. Children needed to infer that—under the task rules that the opponent took the box to which they pointed—they could now win the treat on every trial if they pointed to the empty box. However, nearly all of the 3-year-olds in Russell and colleagues' study failed to do this on the first trial, pointing instead to the box with the treat—and a majority of the children continued to point to (and therefore lose) the treat on all 20 test trials. Similar results have been observed in a number of other studies (e.g., Hala & Russell, 2001; Hughes & Russell, 1993; Russell, 1996; but see Russell, Hala, & Hill, 2003, and Samuels, Brooks, & Frye, 1996).

Although children's difficulty with tasks of this nature appear to be remarkably robust, a number of studies have shown that children's difficulties can be substantially reduced by changing the means by which they give their response. Carlson, Moses, and Hix (1998) used a paradigm somewhat similar to the Windows task in which children were encouraged to indicate the wrong box for an experimenter to open. When children indicated their chosen box by pointing with their hand, they performed relatively poorly (an average of 0.6 correct responses out of 3 attempts). In contrast, when children responded by placing a marker on their chosen box or by pointing to a box using a rotating arrow, their performance was significantly better (an average of 1.8 correct responses out of 3 attempts). Children who responded by pointing with an arrow were significantly more successful at indicating the empty box than children who pointed with their hand, and this benefit was apparent from the very first trial. Performance on the Windows task was also improved when children responded with an arrow or a marker (Hala & Russell, 2001). The authors of these two studies offered very different explanations for how this response mode effect arises. But although this question remains to be resolved, there is agreement on the basic finding that nonstandard response modes improve children's performance. Importantly, although these tasks are often glossed as requiring children to trick an opponent, there is direct evidence that deception in particular, and "theory of mind" in general, is not the source of children's difficulty. Russell and colleagues (1994) found no difference in children's error rates between versions of the Windows task that involved deception and versions that did not, and subsequent studies have confirmed that children have difficulty on the Windows task when there is no opponent to deceive (Carroll, Apperly, & Riggs, 2007a, 2007b). Thus, although Carlson and colleagues (1998) found that an arrow assists children's performance on a test of strategic deception, there are grounds for thinking that the arrow did not help children with the deceptive component of the task.

The beneficial effect of pointing with an arrow rather than the hand has also been reported on tests of counterfactual reasoning; for example, 3-, 4-, and 5-year-olds were significantly more likely to respond correctly to counterfactual conditionals and counterfactual syllogisms (''All fish live in trees. Tom is a fish. Does Tom live in a tree or in water?'') when they indicated their answer by pointing with an arrow (61% correct) than when they pointed with their hand (39% correct) (Beck, Carroll, Brunsdon, & Gryg, 2011). In the counterfactual reasoning tasks, the Windows task, and Carlson and colleagues' (1998) boxes task, children were required to ignore a more salient response in order to indicate a less salient response. Despite the obvious differences between these tasks, in all contexts preschoolers fared significantly better when they gave their response using a nonstandard response mode. Our interest in the current study was to investigate the mechanism of this effect in two ways: first, to establish how far the benefit generalizes to other areas of cognition and, second, to determine what this tells us about cognitive development during the preschool years.

The first question of the current study is a straightforward one—to identify how far the response mode effect extends. We do not yet know whether this response mode benefit is limited to particular tasks or is instead a universal means by which young children can demonstrate previously unrevealed ability. Responding by pointing with the hand may be particularly difficult for young children, and it could be that *any* task requiring a manual point as a response is open to improvement by this means. On the other hand, the response mode effect may be specific only to some as-yet-unidentified common task demand made by the tasks mentioned above. To explain the effect, it is imperative to be able to better identify the situations where the benefit does and does not arise. Experiment 1 addressed this issue directly by comparing different response modes on a selection of preschool cognitive tasks. Our second question, that of explaining how nonstandard response modes help children, was dealt with more fully in Experiment 2.

To inform our choice of tasks, we briefly consider existing accounts offered to explain how nonstandard response modes help children (and, by extension, how older children come to succeed on such tasks spontaneously). The first of these, based on Carlson and colleagues' (1998) study, we call the ''response execution'' account. According to this view, children are unable to respond on the task in the way they intend because they cannot inhibit an incorrect response made prepotent by their previous response histories. Specifically, this account posits that children are accustomed to using a pointing gesture informatively and veridically; to use the same gesture in the unusual way required for success on the task, children must inhibit a tendency to point to where the key object is. This account suggests that despite their persistent failure on strategic reasoning tasks, children are aware that the correct response is to point to the empty box but are unable to execute this. In other words, this account suggests that responding by pointing with an arrow or by placing a small marker helps children at the stage of *carrying out* a response.

In contrast, the ''response formulation'' account suggests that nonstandard response modes provide support at the stage where the task-appropriate response strategy is *formulated.* This account, based on Hala and Russell (2001) and Apperly and Carroll (2009), focuses on the reasoning demands posed by tasks. For example, to succeed on the Windows task, children must integrate the rule they are taught during the training phase (that the box they point to will be given away) with the new information about the treat's location that only becomes available during the testing phase. That is to say, children must go beyond the rules of the game (that the box they point to will be lost) and the information given in the prompt (''Point to the box for [the experimenter] to open'') to infer an appropriate response strategy—something along the lines of ''In order to win the treat, I should point to the empty box.'' This view is consistent with findings from Simpson, Riggs, and Simon (2004). In their ''rule-given'' version of the Windows task, they sought to remove the need to infer a response strategy by making the response contingency explicit, telling children that they would receive a sticker if they pointed to the empty box. Children still needed to execute a pointing response away from the reward object; nevertheless, in this condition, even the youngest children were able to succeed from the very first trial. However, the wording used in Simpson and colleagues' study was such that the task could conceivably be passed by children unreflectively following a simple verbal instruction and entirely ignoring key features of the Windows task. So although this study showed that children are not entirely stimulus bound in their responding, the authors' conclusion—that the poor performance on the task arises from children's difficulty in formulating an appropriate response strategy—remains to be demonstrated conclusively.

Tasks that show a benefit of nonstandard response modes have been claimed to make demands of both children's inhibitory control and their reasoning abilities. Although we acknowledge that the relationship between these two constructs is likely to be complex, it would nevertheless be informative to clarify whether the facilitating effect derives principally from one or both of these abilities. It would also be desirable to be more specific about what aspect of "reasoning abilities" might be implicated in this effect. Reasoning is not an all-or-nothing concept, and the label "reasoning" may refer to quite distinct abilities; reasoning based on conceptual knowledge (such as that involved in mental state understanding) may be very different from reasoning based on an arbitrary task convention (such as on the Windows task). Therefore, in Experiment 1, besides seeking to replicate the positive effect of a nonstandard response mode on children's strategic reasoning on the Windows task (Hala & Russell, 2001), we tested children on tasks that measure inhibitory control and knowledge of mental states.

The response execution account predicts that any task that required children to inhibit a manual pointing response should show a benefit when children are offered an alternative way of responding. The strongest test of this claim would be to use a task that taxed children's inhibitory control without posing any other significant cognitive demand. One such task is the Grass/Snow task (Carlson & Moses, 2001). On this task, children are shown a white square and a green square and are told to point to the white square when the experimenter says "Grass" and to point to the green square when the experimenter says "Snow." Like the Windows task, children must give a counterintuitive pointing response to a two-item array. Unlike the Windows task, however, children are given full explicit instructions for how to respond under all eventualities ("If 'Snow,' then point to green; if 'Grass,' then point to white"). Although children must hold in mind two rules on this task, this is not difficult for them (Simpson & Riggs, 2005). Furthermore, there is no requirement that children draw inferences or go beyond the information presented to them. As such, the Grass/Snow task offers a relatively pure measure of inhibitory control.

In contrast, measures of mental state reasoning rely both on children's conceptual knowledge and on their executive function. For this reason, such tasks are of particular interest in exploring the effect of nonstandard response modes on children's cognition. It is well documented both that inhibitory control plays a role in children's ability to demonstrate their mental state understanding (e.g., Moses, 2001) and that although inhibitory control may be necessary for theory of mind, it is clearly not sufficient (Hughes, 2002; Moses & Tahiroglu, 2011). There remains a significant task demand that cannot be attributed to executive factors and that is best explained by children's ability to make judgments about the mental states of others (Wellman et al., 2001). In other words, although altering the incidental demands of the task can reduce the age at which children first succeed, it cannot remove the requirement that children understand that other people can have mental states different from their own and so should not be expected to remove their difficulties entirely.

If the arrow helps children to inhibit an inappropriate response, then we would expect a significant improvement in performance, similar to previously reported executive manipulations such as the presence or absence of a crucial item in the experimental story (Brown & Bull, 2007; Wimmer & Perner, 1983). To allow a direct test of this idea, we used two variants of the Unexpected Transfer task in Experiment 1. Both variants involve a change of location of an object in order to create a false belief in the main character. In the object-present version, this item will move to a new location in the array. In the object-absent version, the item will be removed from the array completely. The presence or absence of the key object has been shown to directly affect children's performance on the task, with the object-absent version allowing more children to succeed on the task. Removing the object is thought to reduce the need for children to inhibit reference to the current location of the object, meaning that the inhibitory demands of the task are reduced (Flynn, 2007; Moses, 2001). If pointing with an arrow also reduces inhibitory demands, then we should expect to see improvements when (a) the object is present and children point with an arrow and when (ii) the object is absent and children point with their hand.

On the other hand, if the arrow helps children to work out appropriate responses, then the prediction for the Unexpected Transfer task is less obvious. Either this reasoning benefit conveyed by the arrow is sufficiently general that it can help children to demonstrate conceptual knowledge that they do not otherwise possess or the requirement for conceptual knowledge on measures of theory of mind is so great that simply altering the way in which responses are given cannot overcome this. For the moment, we leave further discussion of this question until after Experiment 1.

In the following experiment, we compared two ways of responding on three different tests of preschool cognition. Children used either their index finger or a rotating arrow to respond on the Windows task, the Unexpected Transfer task, and the Grass/Snow task. The three tasks chosen are of particular interest both for their commonalities and their differences. Each requires children to make a pointing response to indicate their response. Each involves a two-item array (either two boxes, two story locations, or two colored cards) in which the correct response is likely to be less salient due to factors within the task. In addition, each task has been shown to be sensitive to developmental change in the 3- and 4-year age range. Notably, however, the tasks differ in key ways. The Unexpected Transfer task requires conceptual knowledge about mental states. The Windows task requires children to work out and then execute a counterintuitive response strategy. The Grass/Snow task requires children to inhibit pointing to the response option most strongly suggested by the prompt on a given trial. Together, they offer a clearer picture of when and how nonstandard response modes help preschoolers.

Experiment 1: Finger versus arrow pointing on the Unexpected Transfer task, Windows task, and Grass/Snow task

Method

Participants

A total of 69 children (33 boys and 36 girls) from nursery schools serving a lower middle-class area of the United Kingdom participated in the experiment. Children were predominantly of Caucasian descent. Data from 9 children were excluded for failing to complete a full set of testing trials ($n = 4$) or failing the reality control question on the Unexpected Transfer task ($n = 5$), leaving a final sample of 60 children. The mean age of the children tested was 3 years 8 months (3; 8; range = 3; 1–4; 5). Children were randomly assigned to one of the two experimental conditions (Finger or Arrow condition, both *n*s = 30). A one-way analysis of variance (ANOVA) found no significant age differences between conditions (mean ages: Finger condition, 3; 8; Arrow condition, 3; 7), $F(1,58) = 0.12$, $p = .91$. To examine the effect of age in subsequent analyses, the two conditions were divided by median split into an older group and a younger group (mean ages: Finger younger group, 3; 4; Finger older group, 4; 0; Arrow younger group, 3; 3; Arrow older group, 3; 11; all *n*s = 15).

Materials

For the Arrow condition, a large cardboard arrow mounted on a square cardboard base was used. For the Unexpected Transfer task, small plastic toys were used to act out the scenarios. For the Windows task, two cardboard boxes (one red and one blue) measuring 4 by 4 by 6 inches were used for the training phase. Two similar boxes with windows cut in the side of them were used for the testing phase. For the Grass/Snow task, white and green square cards were used.

Procedure

Children were tested individually in a room adjacent to the main classroom or in a quiet corner of the main classroom itself. Children were seated across a table from the experimenter and were told that they were going to play some fun games. Tasks were completed in a fully randomized order. For all tasks, children in the Finger condition gave their responses by pointing with their finger, and children in the Arrow condition responded by pointing with a rotating arrow. Children in the Arrow condition were shown the rotating cardboard arrow before testing began. They were shown how to make it rotate, and to ensure that they were comfortable with using the arrow before testing began, they received experience of pointing the arrow at themselves, at the experimenter, and at two salient features of the room (a window and a door).

Unexpected Transfer task. There were four scenarios, each of which involved acting out a brief story using toy dolls and props. The scenarios were based on Wimmer and Perner's (1983) unexpected location false belief task, and experimental scripts can be found in the Appendix. In the two object-absent scenarios, the key object in the story was removed, so that it was no longer present in the array. In the two object-present scenarios, the key object was moved to a different location in the array. Therefore, in all scenarios, the protagonist had a false belief, but in only half of the scenarios was the object actually present in the array before the children. Children were asked the test question, "Can you point to where [the character] thinks [the object] is?" For each scenario, children were also asked a reality control question to check whether they could remember key information about the scenario. This was asked after the false belief test question.

Windows task. The task procedure was based on that used by Carroll and colleagues (2007b). During both the testing phase and the training phase, children indicated their chosen box either by pointing with their finger or by pointing with a rotating arrow.

Training phase. Children were first shown the boxes without windows and were told that the experimenter was going to hide a treat in one of the boxes. They were told that they could choose which box the experimenter would look in. It was explained to them that if the treat was in the experimenter's box, then the experimenter would keep it, but that if it was not, then the children would keep it. Children were then asked to close or cover their eyes while the treat was hidden. They then opened their eyes and were asked to point to the box they wanted the experimenter to open. The experimenter opened the box indicated, and if the treat was in that box, then the experimenter kept it; if it was not, then the second box was opened and the children got to keep it. This was repeated for 5 training trials. On Trial 6, after the indicated box had been opened, children were asked, "So who gets the treat this time?" to ensure that children understood the rules. This questioning was repeated until children had given three consecutive correct responses. All children managed this within 10 trials.

Testing phase. The training boxes were replaced with the boxes with windows. The windows were pointed out to children, and the experimenter commented that that the children would now be able to see where the treat was. As during the training phase, a treat was placed in one of the boxes and the children were asked to indicate a box for the experimenter to open, using the wording, "Point to a box for me to open." This continued for 9 trials. After each trial, the experimenter announced in a neutral tone, "I [You] keep the treat this time" as appropriate. Children used the same response mode during the testing phase as they used during the training phase.

Grass/Snow task. This task procedure was based on the methodology described in Carlson and Moses (2001). After verifying that children were able to name the colors of grass and snow, the experimenter showed the children the white and green cards. Children were told to point to the white card when the experimenter said "Grass" and to point to the green card when the experimenter said "Snow." Children pointed either with their finger or with the rotating arrow according to their condition. The experimenter demonstrated 1 "Grass" trial and 1 "Snow" trial, and then children received 4 practice trials. If they made an error on any of these trials, the experimenter corrected them and reminded them of the rules. Children received 16 test trials, and in all conditions the same prompts were used before each trial—either the word "Grass" or the word "Snow".

Results

Unexpected Transfer task

Table 1 shows the mean numbers of correct answers by response mode and scenario type. A $2 \times 2 \times 2$ ANOVA with response mode (finger vs. arrow) and age group (younger vs. older) as between-participant factors and scenario type (object present vs. object absent) as a within-participant factor found a significant effect of scenario type, $F(1,56) = 107.74$, $p > .001$, $\eta^2 = .658$, but no effect of response mode, $F(1,56) = 0.11$, $p = .74$, $\eta^2 = .002$. Children performed better when the key object on the task was removed than when it was present. There was a significant effect of age group,

Table 1
Correct responses on the Unexpected Transfer task, Windows task, and Grass/Snow task by response mode in Experiment 1.

	Response mode	*n*	UTT object present (out of 2)	UTT object absent (out of 2)	Total UTT (out of 4)	UTT%	WT score (out of 9)	WT%	GS score (out of 16)	GS%
All children	Finger	30	0.47	1.40	1.87 (1.40)	46.8	4.12 (3.02)	46.2	11.97 (3.41)	74.8
	Arrow	30	0.53	1.23	1.76 (1.63)	44.0	7.37 (1.92)	81.9	10.83 (4.20)	67.7
Younger children	Finger	15	0.13	1.20	1.33 (0.99)	33.3	2.67 (2.64)	29.7	10.20 (3.34)	63.8
	Arrow	15	0.33	1.20	1.53 (0.72)	38.3	6.47 (2.23)	71.9	9.73 (3.91)	60.8
Older children	Finger	15	0.80	1.60	2.40 (1.06)	60.0	5.67 (2.69)	63.0	13.73 (2.31)	85.8
	Arrow	15	0.73	1.27	2.00 (1.65)	50.0	8.07 (0.88)	89.7	11.93 (4.18)	74.6

Note. Standard deviations are in parentheses. UTT, Unexpected Transfer task; WT, Windows task; GS, Grass/Snow task.

$F(1,56) = 6.61$, $p = .013$, $\eta^2 = .106$. There was a borderline significant interaction between scenario type and age group, $F(1,56) = 3.64$, $p = .062$, $\eta^2 = .061$. No other interactions were significant.

Windows task

The mean number of treats won by children in each condition is shown in Table 1. A 2 × 2 ANOVA with response mode (finger vs. arrow) and age group (younger vs. older) as between-participant factors found a significant effect of response mode, $F(1,56) = 28.90$, $p < .001$, $\eta^2 = .340$, and a significant effect of age group, $F(1,56) = 15.91$, $p < .001$, $\eta^2 = .221$. There was no significant interaction. Children performed better when pointing with an arrow than when pointing with their finger.

Grass/Snow task

The mean numbers of correct responses by condition is shown in Table 1. A 2 × 2 ANOVA with response mode (finger vs. arrow) and age group (younger vs. older) as between-participant factors found a significant effect of age group, $F(1,56) = 10.00$, $p = .003$, $\eta^2 = .151$. There was no significant effect of response mode, $F(1,56) = 1.56$, $p = .22$, $\eta^2 = .027$, and no significant interaction.

Discussion

Experiment 1 found that using a nonstandard response mode leads to better performance on the Windows task. However, there was no effect of response mode on children's performance on the Unexpected Transfer task or on the Grass/Snow task. The results from the Unexpected Transfer task are particularly striking. The role of inhibitory control on mental state reasoning has been documented extensively (Carlson & Moses, 2001; Moses, 2001; Wellman et al., 2001), and it has been shown to influence the expression of theory of mind understanding in preschool children (e.g., Brown & Bull, 2007). The current results are entirely consistent with these findings. Not only is there a clear age-related improvement on the Unexpected Transfer task, but also there is significantly poorer performance on stories where the key object is moved to a new location rather than removed—a variant with greater inhibitory demands. This effect is somewhat more pronounced in the younger age group, consistent with the idea that in the object-present stories these children's immature inhibitory control means that they are less able to ignore a salient current reality. In contrast, their performance on the object-absent stories is significantly better. This suggests that children in the current sample benefit from manipulations that lower inhibitory demands. However, because changing the response mode makes no difference to children's performance on the Unexpected Transfer task, these data would seem to suggest that the arrow manipulation is not, in fact, reducing the relevant inhibitory demands of the task.

There was a similar lack of response mode benefit on the Grass/Snow task. The task taps children's inhibitory control; children must set up two online responses and then, following a prompt (e.g., "Snow"), must inhibit outputting the response typically associated with the prompt (pointing to the white card) in order to make the response less closely associated with it (pointing to the green card). The lack of response mode effect on the Grass/Snow task suggests that the benefit of nonstandard response modes on the Windows task was not the result of reducing demands on response inhibition. These findings are consistent with other unpublished data from our lab (Carroll, 2005).

Experiment 1 is informative in constraining an explanatory account for the response mode effect. It replicates the response mode benefit seen on tests of strategic reasoning while demonstrating that the effect does not generalize to all preschool pointing tasks. These findings argue against the view that the benefit of a nonstandard response mode is to reduce the inhibitory demands of the task. However, although this would seem to rule out one plausible explanation for the effects of response mode manipulations, it does not yield a positive account of how beneficial effects do arise on some tasks. This was the motivation for Experiment 2.

It is likely that reasoning and executive function share an interdependent relationship, and so conclusively identifying the relative contribution of each on any task is unlikely to be a straightforward process. However, one clear interpretation of the data from Experiment 1 is that changing the response mode allows children to more easily work out what to do. To offer a more direct test of this hypothesis, we ran a second experiment using a further variant of the Windows task. In this experiment, we included the Finger and Arrow conditions of Experiment 1. We also introduced a third condition that we called the Mixed condition, in which children responded with a rotating arrow for the first 3 trials and then responded with their finger on subsequent trials. If the benefit of the rotating arrow is helping children with the reasoning required by the task, then performance in the Mixed condition should remain high even when children are required to point with their finger; their initial experience in pointing with the arrow should help them to work out a task-appropriate strategy for responding that can continue to guide their responding even without the arrow being present. Conversely, if the benefit conveyed helps children with the trial-by-trial executive costs of carrying out a response, then performance should deteriorate when the arrow is withdrawn.

Experiment 2: Investigating further the effect of nonstandard response modes

Method

Participants

A total of 79 children attending three nursery schools serving lower middle-class and working-class areas of the United Kingdom participated (41 boys and 38 girls). Children were predominantly of Asian or Caucasian descent. The mean age of the sample was 3 years 8 months (3;8; range = 3;2–4;3). Children were assigned at random to one of three experimental conditions. A one-way ANOVA confirmed that the mean ages of the three groups were not significantly different (mean ages: Finger, 3;8 [$n = 25$]; Arrow, 3;8 [$n = 29$]; Mixed, 3;9 [$n = 25$]), $F(2,76) = 1.24$, $p = .29$.

Materials

These were identical to the Windows task materials used in Experiment 1.

Procedure

Participants were tested individually either in a quiet corner of the main nursery classroom or in a room or corridor adjacent to the main classroom. The experimenter and children sat facing each other on opposite sides of a table. Children were told that they were going to play a game with the experimenter to see which of them could win more treats. Children underwent a training phase to ensure that they understood the rules of the game as well as the method of responding.

Training phase and testing phase: Finger condition and Arrow condition. These were identical to those in the procedure used in Experiment 1.

Training phase: Mixed condition. This was identical to the training phase for the Arrow condition reported in Experiment 1.

Testing phase: Mixed condition. For the first 3 testing trials, the procedure was identical to that of the Arrow condition. After 3 trials, the experimenter announced, “We don’t need to use this anymore; you can just point with your hand now,” and removed the arrow. For Trials 4 to 9, the procedure was identical to that of the Finger condition.

Results

Overall performance by condition is shown in Table 2. For the following analyses, children were divided into younger and older age groups, calculated by a median split within each condition (mean ages, younger/older: Finger, 3;5/3;11; Arrow, 3;5/3;11; Mixed, 3;6/4;0). A 2 × 3 ANOVA with age group (younger vs. older) and condition as factors found significant effects for both age, $F(1,73) = 8.36$, $p = .005$, $\eta^2 = .10$, and condition, $F(2,73) = 14.63$, $p < .001$, $\eta^2 = .29$. There was no significant interaction. Tukey’s post hoc tests showed that children in both the Arrow condition ($p < .001$) and Mixed condition ($p < .001$) performed significantly better than children in the Finger condition. There was no difference between the Mixed and Arrow conditions ($p = .99$).

Table 2
Correct responses on the Windows task by condition in Experiment 2.

	Response mode	*n*	Trials 1–3	Trials 4–9	Trial 4	Trial 5	Total (out of 9)
All children	Finger	25	1.08 (1.11)	3.60 (2.14)	0.48	0.44	4.68 (3.02)
	Arrow	29	1.97 (1.09)	5.55 (0.69)	0.76	0.90	7.52 (1.62)
	Mixed	25	2.52 (0.77)	5.00 (1.44)	0.60	0.80	7.52 (1.85)
Younger children	Finger	13	0.85 (0.99)	2.77 (2.31)	0.38	0.38	3.62 (3.04)
	Arrow	14	1.50 (1.09)	5.29 (0.83)	0.64	0.86	6.79 (1.72)
	Mixed	12	2.50 (0.91)	4.75 (1.48)	0.50	0.67	7.25 (2.22)
Older children	Finger	12	1.33 (1.23)	4.50 (1.57)	0.58	0.50	5.83 (2.66)
	Arrow	15	2.40 (0.91)	5.80 (0.41)	0.87	0.93	8.20 (1.21)
	Mixed	13	2.54 (0.66)	5.23 (1.42)	0.69	0.92	7.77 (1.48)

Note. Standard deviations are in parentheses.

To investigate how the withdrawal of the facilitating response mode affected children's performance, mean performance across Trials 4 to 9 was compared across all conditions. An ANOVA with age and condition as factors found significant effects for age, $F(1,73) = 7.83$, $p = .007$, $\eta^2 = .10$, and condition, $F(2,73) = 12.24$, $p < .001$, $\eta^2 = .25$. There was no significant interaction. Tukey's post hoc tests showed that children in the Arrow condition ($p < .001$) and Mixed condition ($p = .005$) performed better than children in the Finger condition but that there was no difference between the Arrow and Mixed conditions on Trials 4 to 9 ($p = .55$).

Discussion

As in Experiment 1, children who used an arrow to respond on the Windows task performed significantly better than children who pointed with their hand. Crucially, children in the Mixed condition performed significantly better than those in the Finger condition on Trials 4 to 9 when they needed to respond by using their finger to point. Therefore, even after the arrow had been withdrawn, these children continued to outperform children in the Finger condition despite both groups indicating their responses in exactly the same way. As in previous studies that used comparable mixed conditions, children in the Mixed condition showed signs of a dip in performance on the trial after the arrow was withdrawn, followed by rapid recovery (60% on Trial 4 vs. 80% on Trial 5) (see also Apperly & Carroll, 2009; Carroll et al., 2007a, 2007b). However, performance in the Mixed condition on Trials 4 to 9 was no different from that in the Arrow condition and was significantly better than that in the Finger condition, and this sustained benefit was evident in both the younger and older halves of the sample. This offers clear support for the idea that initially pointing with an arrow allows children to formulate a basis for responding that continues to help even after the response mode reverts to the more challenging pointing with the finger.

General discussion

The experiments reported here present convergent evidence from three quite different tasks exploring how alternative response modes can help children to overcome difficulties associated with manual pointing. Findings from both experiments indicate that this benefit is more likely to arise by supporting aspects of children's reasoning rather than by simply reducing demands on their executive function. Of the three tasks used, the one with the purest demands on children's inhibitory control is the Grass/Snow task. Children's principal difficulty on this task is with resisting the tendency to point toward the response typically associated with the prompt word. We found that children who responded by pointing with an arrow performed no better than children who pointed with their hand. (Further studies conducted in our lab [*N*s = 125 and 57] also found no effect of changing the means of responding on the Grass/Snow task, either for pointing with an arrow or for placing a marker to indicate a response; comparable findings with other nonstandard response modes on this task were reported by Simpson & Riggs, 2009).

There was also no beneficial effect of using a nonstandard response mode on the Unexpected Transfer task. This task assesses children's understanding of mental states while also posing incidental task demands that are executive in nature—notably the need to inhibit any tendency to respond based on children's own perspective. Experiment 1 found a significant effect of removing the key object on the task (thereby reducing inhibitory demands) but no effect of getting children to respond using an arrow. These findings show that the current sample of children was able to benefit from a manipulation that reduced inhibitory demands. It is particularly informative that there was no difference in performance between children who indicated their choice of response by pointing with an arrow and those who pointed with their hand.

The only task where an altered response mode helped children was the measure of strategic reasoning, the Windows task, and findings from Experiment 2 shed further light on the nature of this effect. If the benefit of responding with an arrow was that it helped children to inhibit an incorrect response tendency, then we would have expected children's performance to get worse when the arrow was removed and they were made to point with their hand. Instead, Experiment 2 found that responding with an arrow for 3 trials effected a lasting improvement on children's performance that sustained even when children went onto point with their hand. (This sustained beneficial effect of the arrow has also been replicated in a further study in our lab [$N = 79$].)

Thus, the extent and nature of the response mode benefit have been better identified. What remains to be discussed is the mechanism by which a nonstandard response mode leads to significantly better performance under some circumstances but not others. The absence of any beneficial effect on the Grass/Snow and Unexpected Transfer tasks, coupled with the presence of a sustained and beneficial effect on the Windows task, suggests that nonstandard response modes assist children with a task demand that is absent from the former two tasks but present in the latter task. We note too that Beck and colleagues (2011) reported better performance from 3-, 4-, and 5-year-olds on counterfactual reasoning tasks when children indicated their response by pointing with an arrow rather than with their finger. A parsimonious account of the response mode benefit would identify the task demand present on measures of strategic reasoning and counterfactual reasoning but absent from measures of inhibitory control and mental state understanding. We suggest that this task demand is likely to be closely related to a reasoning component of the tasks given that both strategic reasoning and counterfactual tasks require children to make inferences to reach a task-appropriate conclusion.

If it is indeed the case that nonstandard response modes help children in tasks that require reasoning, then it still remains to be explained how this effect occurs. Our preferred explanation involves an interaction between executive demands (specifically impulse control) and reasoning. One obvious difference between pointing with an arrow and pointing with a finger is that the latter can be done much more quickly. If pointing with a finger enables children to make a fast unreflective action, then their response is likely to be determined by whatever response is currently prepotent rather than one that is the product of any reasoning about the situation. On the Windows task, the overall task goal of wanting to win the treat means that the initial prepotent response for children is likely to be selecting the box that contains the reward item. In contrast, a slower response mode such as a rotating arrow would prevent the rapid execution of a response, helping to avoid impulsivity and opening the way for children to respond on the basis of their reasoning (see also Beck et al., 2011, and Simpson et al., 2012). It has been argued elsewhere (Apperly & Carroll, 2009; Carroll et al., 2007a, 2007b; Simpson et al., 2004) that reasoning is critically necessary to infer the correct response strategy of "pointing to the empty box" and that children should have no difficulty once this strategy has been inferred. This pattern was observed in Experiment 2 of the current study, where on the Windows task children's success on Arrow condition trials was sustained on subsequent trials when they pointed with their finger. This account would suggest that once a successful basis of responding has been established, children can use this to guide their performance on subsequent trials even when more challenging task conditions are reintroduced. In other words, the initial prepotent error of pointing to the baited box is removed thanks to children formulating a new task goal (perhaps along the lines of "Point to the box I don't want"). Even if children then responded impulsively, this would still lead to good performance because the unreflective prepotent response from this task goal would be to select the optimal response option—the empty box. This view has a common cause with other studies noting that goal-setting is a critical aspect of developing executive control during the preschool years (e.g., Blaye & Chevalier, 2011; Towse, Lewis, & Knowles, 2007). The current findings provide important insights into ways in which performance on some tasks may be influenced by children's ability to set the appropriate goals *for themselves*.

At first glance, this analysis might seem to be inconsistent with our finding that children were not helped in their reasoning in a theory of mind task, especially given that the Unexpected Transfer task has been labeled a counterfactual reasoning task by many researchers (e.g., Riggs, Peterson, Robinson, & Mitchell, 1998). However, we suggest that although the two tasks make common reasoning demands, success on an Unexpected Transfer task also requires children to possess some conceptual knowledge about mental states. On this view, although reasoning ability is a necessary condition for children's success on the Unexpected Transfer task, it is not sufficient. If we suppose, instead, that conceptual knowledge about mental states is the limiting condition on children's success on false belief tasks, then there is no reason to view the current findings as surprising because altering the response mode should not alter children's conceptual knowledge and so should not alter their performance on the Unexpected Transfer task.

Altogether, the current findings add to our understanding of the processes that might contribute to the development of goal-oriented behavior in young children. Three types of change are often implicated in the development of cognitive control in preschoolers. First, there is evidence for quantitative improvement in the capacity for inhibitory control and working memory (Carlson & Moses, 2001; Carlson, Moses, & Breton, 2002). Second, it has been suggested that children's cognitive flexibility is transformed by their ability to hold in mind increasingly complex *structures* that allow the representation of hierarchical rules (Zelazo & Frye, 1997; Zelazo, Mueller, Frye, & Marcovitch, 2003). Third, it has been suggested that children develop increasingly abstract "symbolic" forms of mental representation, making it possible to think about a problem without being distracted by salient characteristics of the stimuli (Apperly & Carroll, 2009; Carlson, Davis, & Leach, 2005; Mischel et al., 1989). In addition to this list, we suggest that it is important to consider *strategic reformulation* as a way in which children might change a task from one beyond the limits of their cognitive resources (e.g., memory and inhibitory control) to one that is less demanding. The current data suggest that such a reformulation process may be induced in preschoolers by manipulations that diminish impulsive responding in children to allow the expression of products of a reasoning process. It is for future work to determine whether some of the developmental changes in flexible goal-oriented behavior in preschoolers are due to children improving their abilities to effect strategic reformulation for themselves.

Acknowledgments

The authors thank Anna Bryant for her help in piloting an earlier version of Experiment 1 and also thank the children, parents, and staff of the participating schools.

Appendix. Unexpected Transfer task scenarios

At the start of the testing session, the experimenter said, "Now I'm going to tell you some short stories".

Story A

Kevin is in the kitchen making a sandwich. Here he is. He decides to put some cheese in it, and some tomato, and some lettuce. There, it's all ready. Before he has time to eat the sandwich, he remembers that he has forgotten to tidy up his bedroom. So he decides to eat the sandwich later after he has tidied up his bedroom. He puts his sandwich in his lunchbox, and off he goes to tidy his room. While Kevin is tidying his room, his little sister Olivia comes along. She is feeling very hungry. She sees the sandwich in the lunchbox and decides to eat it. She is so hungry, she eats the sandwich all up. It's all gone! Then Olivia goes outside to play in the garden. Now Kevin has finished tidying his room. He wants to eat his sandwich".

Experimental question

Can you point to where Kevin thinks the sandwich is?

Check question

Where is the sandwich really?

Story B

"Sandra has got a new bag. She's very pleased with her bag because she thinks it's a lovely color. She doesn't want to lose it, so she decides she'll put it away in the blue toybox. Then she goes upstairs to do some coloring. While she's doing that, her brother Joe comes by. He sees her new bag in the blue toybox and decides to take it out to have a closer look at it. He also thinks it's a very nice bag. Then he decides to put it away safely in the red toybox and goes to watch TV. Sandra has finished her coloring and wants to put the picture she has done into her bag".

Experimental question

Can you point to where Sandra thinks the bag is?

Check question

Where is the bag really?

Story C

"Andrew is going to a party and decides he will dress up very smart. He puts on his suit and then puts on his nicest hat. He does look smart! But then he sees that he's forgotten to brush his hair. So he takes off his hat and puts it in his hat box and goes upstairs to brush his hair. His friend Suzie comes in. She likes Andrew's hat and thinks she'll try it onto see what she looks like. So she takes it out of the hatbox and tries it on. She thinks it doesn't really suit her, so she puts it away in the red toybox and goes outside to play. Andrew finishes brushing his hair and wants to put on his hat again".

Experimental question

Can you point to where Andrew thinks the hat is?

Check question

Where is the hat really?

Story D

"Jane has just come home from the shops. She's bought herself an apple to eat. She's just about to eat it when the telephone rings. She puts her apple in the red toybox and goes to answer the phone. While she's talking on the telephone, her friend Simon comes in. He sees the apple and thinks that Jane must have bought it for him, so he eats it all up. It was delicious! Then he goes to his bedroom to read a book. Jane has finished talking on the phone. Now she wants to eat her apple".

Experimental question

Can you point to where Jane thinks the apple is?

Check question

Where is the apple really?

References

Apperly, I. A., & Carroll, D. J. (2009). How do symbols affect 3- to 4-year-olds' executive function? Evidence from a reverse-contingency task. *Developmental Science, 12*, 1070–1082.

Beck, S. R., Carroll, D. J., Brunsdon, V. E. A., & Gryg, C. K. (2011). Supporting children's counterfactual thinking with alternative modes of responding. *Journal of Experimental Child Psychology, 108*, 190–202.

Blaye, A., & Chevalier, N. (2011). The role of goal representation in preschoolers' flexibility and inhibition. *Journal of Experimental Child Psychology, 108*, 469–483.

Brown, E. L., & Bull, R. (2007). Can task modifications influence children's performance on false belief tasks? *European Journal of Developmental Psychology, 4*, 273–292.

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

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*, 1032–1053.

Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relationship between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development, 11*, 73–92.

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*, 672–691.

Carroll, D. J. (2005). Interactions between children's reasoning and executive function. Unpublished doctoral dissertation, University of Birmingham.

Carroll, D. J., Apperly, I. A., & Riggs, K. J. (2007a). Choosing between two objects reduces 3-year-olds' errors on a reverse-contingency test of executive function. *Journal of Experimental Child Psychology, 98*, 184–192.

Carroll, D. J., Apperly, I. A., & Riggs, K. J. (2007b). 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*, 142–148.

Flynn, E. (2007). The role of inhibitory control in false belief understanding. *Infant and Child Development, 16*, 53–69.

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

Hughes, C. (2002). Executive functions and development: Emerging themes. *Infant and Child Development, 11*, 201–209.

Hughes, C., & Russell, J. (1993). Autistic children's difficulty with mental disengagement from an object. *Developmental Psychology, 29*, 498–510.

Mischel, W., Shoda, Y., & Rodriguez, M. (1989). Delay of gratification in children. *Science, 244*, 933–938.

Moore, C., Barresi, J., & Thompson, C. (1998). The cognitive basis of prosocial behavior. *Social Development, 7*, 198–218.

Moses, L. J. (2001). Executive accounts of theory of mind development. *Child Development, 72*, 688–690.

Moses, L. J., & Tahiroglu, D. (2011). Clarifying the relation between executive function and children's theories of mind. In B. W. Sokol, U. Müller, J. I. M. Carpendale, A. R. Young, & G. Iarocci (Eds.), *Self and social regulation: Social interaction and the development of social understanding and executive functions* (pp. 218–233). Oxford, UK: Oxford University Press.

Riggs, K. J., Peterson, D. M., Robinson, E. J., & Mitchell, P. (1998). Are errors in false belief tasks symptomatic of a broader difficulty with counterfactuality? *Cognitive Development, 13*, 73–90.

Russell, J. (1996). *Agency: Its role in mental development*. Hove, UK: Lawrence Erlbaum.

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*, 111–137.

Russell, J., Jarrold, C., & Potel, D. (1994). What makes strategic deception difficult for children—The deception or the strategy? *British Journal of Developmental Psychology, 12*, 301–314.

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 through the Windows task. *British Journal of Developmental Psychology, 14*, 159–172.

Simpson, A., & Riggs, K. J. (2005). Inhibitory and working memory demands of the Day–Night task in children. *British Journal of Developmental Psychology, 23*, 471–486.

Simpson, A., & Riggs, K. J. (2009). What makes responses prepotent for young children? Insights from the Grass/Snow task. *Infant and Child Development, 18*, 21–35.

Simpson, A., Riggs, K. J., Beck, S. R., Gorniak, S. L., Wu, Y., Abbott, D., et al (2012). Refining the understanding of inhibitory processes: How response prepotency is created and overcome. *Developmental Science, 15*, 62–73.

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*, 155–170.

Sodian, B. (1994). Early deception and the conceptual continuity claim. In C. Lewis & P. Mitchell (Eds.), *Children's early understanding of mind: Origins and development* (pp. 385–401). Exeter, UK: Lawrence Erlbaum.

Towse, J. N., Lewis, C., & Knowles, M. (2007). When knowledge is not enough: The phenomenon of goal neglect in preschool children. *Journal of Experimental Child Psychology, 96*, 320–332.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72*, 655–684.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*, 103–128.

Zelazo, P. D., Mueller, U., Frye, D., Marcovitch, S. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development*, 68 (3, Serial No. 274).

Zelazo, P. D., & Frye, D. (1997). Cognitive complexity and control: A theory of the development of deliberate reasoning and intentional action. In M. Stamenov (Ed.), *Language structure, discourse, and the access to consciousness* (pp. 113–153). Amsterdam: John Benjamins.