The Control of Instrumental Action Following Outcome Devaluation in Young Children Aged Between 1 and 4 Years

U. M. H. Klossek, J. Russell, and A. Dickinson
University of Cambridge

To determine the role of action–outcome learning in the control of young children’s instrumental behavior, the authors trained 18- to 48-month-olds to manipulate visual icons on a touch-sensitive display to obtain different types of video clips as outcomes. Subsequently, one of the outcomes was devalued by repeated exposure, and children’s propensity to perform the trained actions was tested in extinction. On test, children with a mean age greater than 2.5 years performed the action trained with the devalued outcome less than those trained with the still-valued outcome, thereby demonstrating that their actions were mediated by action–outcome learning. By contrast, the instrumental responses of younger children (mean age <2 years) were resistant to outcome devaluation and may have been elicited directly by the icons associated with each response, rather than mediated by a specific action–outcome expectation.

**Keywords:** goal-directed action, instrumental learning, outcome devaluation, children

Most human adults rarely regard the capacity for intentional, goal-directed instrumental action as a remarkable achievement. We do it all the time: We are able to perform an action in the belief that it will bring about a goal or a change in the world that fulfills our needs and desires. Critically, a course of intentional action can be selected and initiated even though at the time of performing the action, the goal is not physically present. By being able to rely on internal representations of the desired goal or outcome and of the causal action–outcome relationship, intentional, goal-directed action enables agents to transcend their current physical environment and extend their sphere of influence to objects and events beyond the immediate present and local space.

Although the capacity for goal-directed action in normal human adults does not require empirical demonstration, the intentional status of young children’s instrumental actions is less clear. Studies of instrumental learning in infancy have shown that infants, and even newborns, are sensitive to the contingent consequences of their behavior (Kalnis & Bruner, 1973; Rovee & Rovee, 1969; Siqueland, 1969; see also Hauf, Elsner, & Aschersleben, 2004) and that they reliably acquire a new instrumental action when it immediately causes an attractive outcome, but not when an equivalent outcome is presented independently or noncontingently (e.g., Rovee-Collier, Morrongiello, Aron, & Kupersmidt, 1978; Siqueland & DeLucia, 1969). This learning was remarkably robust, as the novel response was retained for as long as 5–6 days by 2- to 3-month-olds in the absence of any further training or exposure to the training situation (Rovee-Collier, 1983). However, most standard instrumental learning paradigms are designed to assess learning simply in terms of response acquisition. Therefore, although demonstrating that infants possess remarkable learning capacities even in the first few months of life, these studies do not permit inferences to be drawn regarding the intentional status of these actions. Establishing the goal-directed status of an action requires the demonstration that it is performed specifically to bring about a particular outcome, rather than just that it is acquired when it is followed by an attractive effect.

Determining the intentional status of an instrumental action in the absence of a reliable verbal report from the agent is, however, not an easy matter. To address this issue empirically, students of animal learning have developed a behavioral assay for intentional, goal-directed action by devaluing the outcome following instrumental training. The concept of goal-directedness assayed by this procedure is one in which the action is goal-directed if performance is mediated by the interaction between two representations (Dickinson, 1985, 1989): The first is a representation of the causal or contingent relationship between the action and the outcome, whereas the second is a representation of the current value of the outcome. It is mediation by the action–outcome representation that ensures that the action is directed toward obtaining the outcome, whereas the dependence on a representation of the outcome value ensures that the outcome functions as a goal in controlling performance.

In the prototypical outcome devaluation experiment (e.g., Adams & Dickinson, 1981), animals are initially trained to perform an instrumental action to gain access to a particular outcome, which is subsequently devalued. Although the effectiveness of the devaluation treatment depends on the nature of the outcome, a common procedure is the induction of specific satiety for the outcome through simple exposure (Balleine & Dickinson, 1998b; Colwill & Rescorla, 1985). This devaluation treatment must occur
in the absence of the opportunity to perform the instrumental response to prevent the formation of any direct association between the instrumental action and the devaluation treatment. Finally, the impact of the outcome devaluation is assessed through the subsequent performance of the instrumental action. If the action is goal-directed, performance of the action trained with the now devalued outcome should be reduced in this test relative to a control condition in which the outcome is not devalued. The critical feature of this test is that postdevaluation performance is measured in extinction and therefore in absence of the outcome. Testing in extinction ensures that performance is mediated by the representations the instrumental contingency established during training rather than through the direct impact of the devalued outcome during the test.

We applied the same logic to investigate whether young children represent their own instrumental actions in terms of specific action–outcome relationships. There is currently little evidence that young children are able to select a specific action on the basis of the current value or utility of the past outcome of the action. In fact, studies that examined problem solving or planning, using, for example, child-appropriate versions of standard adult planning tasks, such as the Tower-of-Hanoi disc-transfer puzzle (Simon, 1975), have often found that 2- to 3-year-olds perform relatively poorly (e.g., Klahr & Robinson, 1981; Kopp, O’Connor, & Finger, 1975; O’Sullivan, Mitchell, & Daehler, 2001; Welsh, 1991). However, these tasks were designed to assess relatively complex cognitive skills—such as reasoning, planning, and rule-guided behavior—and it remains unclear whether young children experience difficulty in acquiring and performing the more basic form of goal-directed action assayed by the outcome devaluation procedure.

Moreover, there are reasons for believing that infants and young children might have difficulty selecting an action on the basis of the current value of the outcome. For example, they perform reaching responses and other actions that supposedly reflect goal-directedness, problem solving, and means–end planning, such as pulling a cloth to retrieve an out-of-reach toy, even when there is no goal, that is, no toy to be retrieved (e.g., Smith, Thelen, Titzer, & McLin, 1999; Willatts, 1999). Young children also have the tendency to reach perseveratively in object retrieval tasks and frequently repeat a learned response even though the action continues to be ineffective in obtaining a desirable outcome (e.g., Fox, Kagan, & Weiskopf, 1979; Gratch & Landers, 1971; Schutte & Spencer, 2002; Spencer, Smith, & Thelen, 2001).

The purpose of the present series of experiments was therefore to determine whether instrumental actions performed by young children are sensitive to devaluation of the outcome. Experiment 1 demonstrated that children are sensitive to the instrumental contingency between the target response (touching an icon on a touch sensitive screen) and the outcome (the presentation of a short video clip). Experiments 2 and 3 then investigated whether devaluing the outcome by specific satiety had an effect on the subsequent instrumental behavior of children ranging between 18 and 48 months of age. Finally, Experiment 4 confirmed that the sensitivity to outcome devaluation was mediated by the instrumental contingency between action and outcome rather than being due to the predictive significance of the icons to which the responses were directed.

Experiment 1

To ascertain that the touch response could be brought under instrumental control, we trained children of the youngest age used in our experiments to reach out and touch two different screen areas marked by red and green butterfly icons. During acquisition, one of the actions was followed by short video clips featuring colorful animated scenes, whereas the other produced no outcome. Given that the video clips were effective rewards for the children, the children should have learned to perform the rewarded action in preference to the nonrewarded one if they were sensitive to the action–outcome contingency.

The second phase assessed the effect of reversing the instrumental contingency so that the action rewarded during acquisition no longer produced the outcome, whereas the previously unrewarded response did. It was important to determine whether the target actions were sensitive to changes in the action–outcome contingency. Unless the children adapted to the shift in contingency by reversing their preference for the two actions, it was possible that response perseveration would have masked any sensitivity to outcome devaluation.

Method

Participants

Participants were 5 boys and 3 girls, aged between 19 and 26 months, with a mean age of 22.8 months (SD = 2.5), who were recruited from independent day care centers. Four children did not complete the session because they became distracted (2) or because of interference by other children (2) and were therefore replaced.

Apparatus and Stimuli

The experiment was run on a Sony Vaio laptop computer (F808K and GRX315MP) connected to a stand-alone flat panel Taxan CV600 LCD monitor with an effective display area of 21.5 × 16 cm and a screen resolution of 640 × 480, which was equipped with a Microtouch capacitive touch-screen system. Two external computer speakers (HK 195) connected to the laptop were situated behind the touch screen. The software used for controlling stimulus presentation and recording of responses was written and compiled using Microsoft Visual Basic Professional 6.0. Two 9 × 7.5-cm icons of a red and green butterfly on a white background were displayed on the screen, one on the left side and the other on the right side. These icons acted as targets for the touch response. Eight short video clips from a children’s cartoon were used as the outcomes. During pretraining, four 10-s clips were used, whereas four different 4-s clips constituted the outcomes during acquisition and reversal. All clips were in color and included music and other sounds but had no explicit verbal content.

Procedure

In this and all of the following experiments, the children were tested while seated at a table within easy reach of the touch-sensitive monitor in a familiar room in their regular day nursery. They were introduced to the apparatus and the investigator by a member of the nursery staff well-known to the child. The child’s
Perform the correct rewarded response relative to the incorrect, was a significant increase over trials in children's tendency to (1, 42). During acquisition, revealed a significant linear trend for response reciprocals both phase, respectively. The reliability of the effects in this and all subsequent analyses was assessed against a Type I error rate of .05.

The analysis revealed no significant effects of the response variable, and therefore we collapsed the discrimination ratios across this variable for presentation. As Figure 1 illustrates, the relative performance of the correct, rewarded response increased over trials, both during acquisition and following reversal of the original response–outcome relationships on Trial 9. This description was confirmed by the statistical analyses. The significant main effect of trial, F(1, 7) = 3.82, MSE = 0.08, did not interact with phase, F(7, 42) = 1.09, p = .386, MSE = 0.01. Separate analyses for linear trend using the coefficients −7, −5, −3, −1, 1, 3, 5, 7 revealed a significant linear trend for response reciprocals both during acquisition, F(1, 42) = 4.41, MSE = 0.49, and reversal, F(1, 42) = 4.16, MSE = 0.12, thereby demonstrating that there was a significant increase over trials in children's tendency to perform the correct rewarded response relative to the incorrect, unrewarded response.

Results and Discussion

Acquisition and reversal performance was assessed by a discrimination ratio in the form of the reciprocal of the total number of responses per trial. Because the first correct response always terminated a trial, a ratio of 1 represents perfect discrimination, with lower values indicating poorer discrimination. The ratios were analyzed by a mixed analysis of variance (ANOVA), with the between-children response variable differentiating performance when right or left responses were correct during acquisition. The within-child variables of phase and trial distinguished performance during acquisition and reversal and on the different trials of each phase, respectively. The reliability of the effects in this and all subsequent analyses was assessed against a Type I error rate of .05.

In summary, children of the youngest age used in the present series of experiments showed an increasing and significant tendency to perform the action that caused the video outcome to occur. Moreover, when the action–outcome assignments were changed, the children reversed their responding accordingly. This experiment has therefore shown that video outcomes are effective rewards for children aged between approximately 1 and 2 years and that their touch responses are sensitive to the instrumental response–reward contingencies. After establishing these necessary prerequisites, the next experiment was designed to examine whether these instrumental actions are sensitive to outcome devaluation.

Experiment 2

In the second experiment, children were initially trained to touch both colored butterfly icons. Touching one butterfly produced video clips from one cartoon series as the outcome, whereas touching the other icon produced clips from a second cartoon series. Following this instrumental training, one of the video outcomes was devalued by presenting it repeatedly to reduce its incentive value by inducing satiety for this video. To assess the effectiveness of this devaluation treatment, we gave half of the children a reacquisition test in which they were given the opportunity to obtain further video outcomes under conditions identical to training. If the incentive or goal value of the video outcome had been reduced by the devaluation procedure, the children should have responded less for this devalued video than for the other outcome, which should have retained its incentive value.

The critical condition to assess whether children’s instrumental responses were goal-directed was the postdevaluation extinction test experienced by the remaining children. During the extinction test, children were also given the opportunity to perform the two instrumental responses on the same display that was used during training. In this condition, however, no outcomes were presented to ensure that children’s performance reflected the knowledge that had been acquired during instrumental training. If performance was mediated by knowledge about the specific response–outcome relationships experienced during training, and if children were able...
to integrate this knowledge with the information concerning the changes in outcome value, they should have performed the response trained with the devalued outcome less than the response that had been trained with the still valued outcome.

To assess whether sensitivity to outcome devaluation varied with age, we recruited the children into three 10-month age bands between 16 and 48 months.

Method

Participants

Seventy-two children aged between 17.7 and 46.7 months were allocated to the extinction condition (18 girls, 18 boys) and the reacquisition condition (17 girls, 19 boys). Within each condition, participants were divided into three 10-month age bands, each comprising approximately equal numbers of boys and girls (see Table 1): 16 to 26.9 months (Age Group 1), 27 to 36.9 months (Age Group 2), and 37 to 48 months (Age Group 3). An additional 43 children (10 in Age Group 1, 17 in Age Group 2, and 16 in Age Group 3) did not complete a full session and were replaced. The reasons were the following: interference by other children, the child’s carer, or other events (17); technical problems or experimenter error (13); spontaneously abandoning the task because of distraction (11); failure to acquire the required touch responses (2).

Apparatus and Stimuli

The touch-screen apparatus and butterfly icons were the same as those used in Experiment 1. Three 24-s video clips from each of two different and highly discriminable children’s cartoons were used as outcomes. All clips were in color and included music and sounds produced by the characters, but had no explicit verbal content.

Procedure

All except 2 of the children were tested at their regular day nursery in the same way as described in Experiment 1. The other 2 children completed the task at our departmental testing facility accompanied by their mothers. Unless stated otherwise, the general procedure was the same as in Experiment 1.

Instrumental training. During the 9-min instrumental training phase, the two butterfly icons were shown side by side in the center of the touch screen display. Initially, the investigator directed the child’s attention to the screen display and then demonstrated each touch response once, thereby triggering the appropriate video presentation. Thereafter, the child was encouraged to touch the pictures and received verbal encouragement and praise for doing so. Every touch on one of the pictures always triggered a clip from the same video series so that one response was always followed by a clip from one of the two cartoons and the other response by a clip from the other cartoon. The assignment of the cartoons to the responses was counterbalanced across children within each group. Simultaneous responses on both pictures did not trigger a video, and children were only allowed to use one hand at a time to touch the pictures. Successive touch responses on the same icon triggered the three successive clips from the appropriate cartoon in a fixed order, regardless of whether clips from the other cartoon were produced in between.

Outcome devaluation. Following training, the butterfly icons disappeared and the three 24-s video clips from one of the two cartoons were repeated four times, each with an interval of 3 s between clips. For half the children, these clips were from one cartoon, with the remaining children being shown the clips from the other cartoon. Because response–outcome assignments and icon location were counterbalanced across participants, this also meant that for half of the children in each group, it was the video outcome associated with left responses that was devalued, and for the remaining half the video outcome associated with right responses was devalued.

Extinction and reacquisition tests. During the tests, the children again had the same opportunity to touch the red and green butterfly icons as during instrumental training. For children in the extinction test, none of their responses triggered any video presentations for a period of 2 min. Thereafter, the icons were activated again for a final 4-min period, which was not part of the test but ended the session on a positive note by allowing the children once again to produce the videos.

For the children in the reacquisition condition, touching the butterfly icons triggered video presentations in the same way as during training. To compensate for the time taken up by the video presentations, we designed the instrumental test in the reacquisition condition to last 6 min.

Results and Discussion

Instrumental Training

The number of responses per minute was calculated for the time during which the children had opportunity to respond to the butterfly icons. Because the variance of the response rates increased with the mean, a square-root transformation was applied to these rates in this and all subsequent analyses before they were evaluated by a mixed ANOVA. The between-children variables were age and condition, which distinguished between performance of the children in the extinction and reacquisition conditions. The within-child variable of devaluation contrasted performance of the action trained with the video outcome that was subsequently devalued with performance of the action trained with the outcome that was still valued at the time of testing.

As Table 2 illustrates, the mean rate of responding increased systematically with age, $F(2, 60) = 14.0, \text{MSE} = 0.81$. More
importantly, however, there was no evidence that the response rates for the actions trained with the “to-be-devalued” and valued outcomes differed reliably during training. There was no significant main effect of devaluation, nor did this variable enter into any significant interactions ($F$s/1).  

**Extinction Test**

Because the response rates during instrumental training varied as a function of age, the rates during the tests were expressed as a percentage of the absolute training rate to minimize the contribution of between-children variance. In the extinction condition, omission of the video outcome on test produced rapid extinction of responding, and some children failed to respond at all in the second minute of the extinction test. Therefore, test rates are based on responding during the first minute of the extinction test. The top panel of Figure 2, which presents responding in the extinction test as a percentage of the training rate, shows that a substantial devaluation effect was observed in the two older groups in that these children performed the response trained with the now devalued outcome at a lower relative rate than the response trained with the valued outcome. By contrast, there was no evidence that the youngest children were sensitive to outcome devaluation, as the relative rates at which they performed valued and devalued responses were not markedly different.

This description was confirmed by statistical analysis. Because the variance of the percentage responding during the tests increased with the mean (see Figure 2), this measure was square-root transformed prior to analysis in this and the subsequent analyses (see Table 3 for the transformed data). There was a significant Age × Devaluation interaction, $F(1, 30) = 4.8, MSE = 27.1$, and analyses of the simple main effects showed that the children performed the valued response more than the devalued one in Age Group 2, $F(1, 11) = 12.0, MSE = 25.4$, and Age Group 3, $F(1, 11) = 19.5, MSE = 25.5$, but not in Age Group 1 ($F < 1$). Furthermore, evidence that the magnitude of the devaluation effect varied with chronological age comes from the significant positive correlation between age and discrimination ratio, where the latter represents the number of valued responses as a ratio of the total number of responses ($r = .36, p = .033$).

**Reacquisition Test**

The pattern of responding during the test in the reacquisition condition shows that the absence of a devaluation effect in the extinction test for the youngest children was not due to a failure of the devaluation treatment to reduce the incentive value of the exposed video clips. As the bottom panel of Figure 2 and Table 3
During the Tests

Table 3  
Mean of the Square-Root Transformed Percentage Responding During the Tests

<table>
<thead>
<tr>
<th>Variable</th>
<th>Devalued response</th>
<th>Valued response</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Experiment 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extinction condition</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age Group 1</td>
<td>9.4</td>
<td>5.2</td>
</tr>
<tr>
<td>Age Group 2</td>
<td>6.5</td>
<td>5.4</td>
</tr>
<tr>
<td>Age Group 3</td>
<td>6.1</td>
<td>3.1</td>
</tr>
<tr>
<td>Reacquisition condition</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age Group 1</td>
<td>9.9</td>
<td>3.9</td>
</tr>
<tr>
<td>Age Group 2</td>
<td>8.1</td>
<td>3.2</td>
</tr>
<tr>
<td>Age Group 3</td>
<td>4.7</td>
<td>2.8</td>
</tr>
<tr>
<td>Experiment 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age Group 1</td>
<td>9.3</td>
<td>6.4</td>
</tr>
<tr>
<td>Experiment 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age Group 2</td>
<td>6.9</td>
<td>4.4</td>
</tr>
<tr>
<td>Age Group 3</td>
<td>6.6</td>
<td>4.7</td>
</tr>
</tbody>
</table>

illustrate, when the video outcomes were presented contingent on responding during reacquisition, children of all ages performed the response that produced the devalued video less often than the one which produced the valued video, $F(1, 30) = 22.9, MSE = 30.6$. Moreover, the fact that the $F$ ratio for the interaction between age and devaluation was less than 1 indicates that the magnitude of the devaluation effect did not vary with age. There was also a significant effect of devaluation in each age group, $F$s($1, 11) > 4.9$, and in contrast to the extinction condition, there was no significant correlation between age and relative performance of the valued response during the reacquisition test ($r = .21, p = .22$).

In summary, the instrumental responses of children aged between 2 and 4 years appeared to be goal-directed. The significant devaluation effect observed in Age Groups 2 and 3 demonstrated that these children could integrate instrumental knowledge acquired during training with information about a subsequent change in the value of one of the outcomes and adapt their behavior appropriately.

Unexpectedly, however, the present results did not provide evidence for goal-directed action in the 18- to 27-month-olds, as the children in this age group responded equally for valued and devalued outcomes in the extinction test. The performance of the youngest reacquisition group confirmed that the absence of a significant devaluation effect in the extinction condition was not due to a failure of youngest children to discriminate between the different types of video outcome, nor a result of the ineffectiveness of the devaluation procedure. When allowed to earn video outcomes directly in the reacquisition test, the youngest children did show a significant preference for the response whose outcome had not been devalued, indicating that the repeated presentations of the relevant video series prior to the test had been successful in reducing the incentive or goal value of the video outcome for this age group. Experiment 3 examined a possible reason why our devaluation procedure may have failed to detect the goal-directed nature of the actions of the youngest children.

Experiment 3

In a series of studies, Rovee-Collier and colleagues (for a review, see Rovee-Collier, 1997) have shown that a failure of instrumental performance by young infants does not always reflect an absence of learning because the presentation of a reminder cue in the form of stimuli from the training context could reinstate performance. The presence of the reminder cue aided retrieval of the training memory and the deployment of this memory in the control of responding. Therefore, it is possible that the youngest children experienced difficulty in retrieving information about the outcome devaluation experience during the extinction test. The fact that the youngest children showed a substantial devaluation effect in the reacquisition condition suggests that actually presenting the video outcome during the test may have aided retrieval of information of the relative values of the two outcomes, thereby allowing the youngest children to manifest differential responding in accordance with outcome value. Therefore, we examined the effect of providing children in the youngest age group with a reminder cue in the form of a still picture from either the devalued or valued video during the extinction test in the hope that this cue would facilitate retrieval of information about the relative values of the two outcomes. If the absence of a devaluation effect in Experiment 2 was due to the fact that the youngest children had difficulty in retrieving this information, then the children should have been more likely to perform the devalued response less than the valued one in an extinction test conducted in the presence of a reminder cue.

Method

Participants, Apparatus, and Stimuli

Twenty-four children (13 boys and 11 girls) matching the age profile of the youngest age group in Experiment 2 (see Table 1) were trained and tested with the same apparatus, stimuli, and video outcomes as in Experiment 2. An additional 16 children did not complete a full session and were replaced for the following reasons: abandoning the task spontaneously (5); fussing or crying (4); equipment malfunction or experimenter error (3); interference by other children, caregivers, or other events (2); failure to perform the required touch responses during instrumental training (2).

Procedure

The instrumental training and outcome devaluation procedures were the same as those used in Experiment 2. The children were trained to touch the two butterfly icons in order to be presented with different video outcomes, before the video that acted as the outcome for one of the responses was devalued by repeated exposure. The children then received an extinction test that was also identical to that in Experiment 2, except for the presence of a reminder cue. This 5 × 7-cm cue was a still picture taken from one of the video cartoons and located above the two butterfly pictures in the center of the screen. For half of the children, the cue was from the devalued video cartoon; the remaining half was presented with a cue from the valued cartoon. Touching the cue picture during the extinction test had no programmed consequences.
Results and Discussion

As the response rates during instrumental training did not differ significantly for the children presented with the valued and devalued reminder cues during extinction testing ($F < 1$), the rates were collapsed for presentations. As Table 2 illustrates, these rates were very similar to those produced by the youngest children in the extinction and reacquisition conditions of Experiment 2, and indeed, the $F$ ratio for the contrast between the response rates of children in the present experiment versus the youngest groups in Experiment 2 was less than 1. As in Experiment 2, the two responses were performed at the same rate during training ($F < 1$).

During the extinction test, the presence of the reminder cues in the present experiment did not induce a devaluation effect. Neither the main effect of the reminder cue valence on the square-root transform of the percentage responding relative to the training rates ($F < 1$) nor the Cue Valence × Devaluation interaction, $F(1, 20) = 1.1, p = .31$, $MSE = 37.7$, was significant, so the test performance was collapsed across the valence of the reminder cue for presentation. The mean percentage response rates for the valued and devalued responses were 125.8 ($SE = 35.4$) and 158.8 ($SE = 39.6$), respectively; which, importantly, did not differ significantly following the square-root transformation ($F < 1$; see Table 3).

The absence of a reliable devaluation effect was not due to the failure of the valence of the reminder cues to transfer to the extinction test. During this test, the children touched not only the target butterfly icons but also the reminder cues and did so more frequently in the case of the valued cue compared with the devalued cue, $t(1, 22) = 2.17$. The mean number of contacts with the devalued cue was 1.0 ($SE = 0.5$), but 3.5 ($SE = 1.0$) with the valued cue. Seven of the 12 children presented with the devalued cue did not touch it, whereas all but 1 of the 12 children touched the valued cue at some point during the test.

Given these data, we think that it is unlikely that absence of a devaluation effect for the youngest children was due to a failure to retrieve information about the relative values of the two outcomes during the test. Rather, it is more likely that the process controlling their touch responding was insensitive to representations of the current values of the outcomes. As mediation by a representation of outcome value is one of our criteria for goal-directedness (see the introduction), responding by the youngest children was not goal-directed by this criterion. We consider the nature of this process in the General Discussion. However, the differential responding to the valued and devalued reminder cues not only strengthens the interpretation of the insensitivity of the youngest children to outcome devaluation but also raises interpretative issues for the outcome devaluation effect observed in the older children.

Experiment 4

The key finding of Experiment 2 was the significant devaluation effect observed in the two older age groups, which suggests that the children of approximately 32 months and above were capable of goal-directed action that was based on knowledge of the instrumental action–outcome contingency. Strictly speaking, however, the procedure employed in Experiment 2 does not warrant this conclusion. The fact that children in Experiment 3 spontaneously touched the valued reminder cue without explicit training or instruction demonstrates that attractive visual stimuli can directly elicit touch responses. At issue, therefore, is whether the children touched the butterfly icons because they learned that these responses produced valued outcomes or because the icons themselves became attractive stimuli through their association with the video outcomes and, as a result, became capable of directly eliciting the response. In other words, the question is whether icon touching was an instrumental response generated by the specific action–outcome contingency or an elicited response generated by the predictive relationship between the butterfly icons and their associated outcomes. By our criteria for goal-directedness (Dickinson, 1985, 1989; Dickinson & Balleine, 1993), only in the former case would we characterize the touch response as goal-directed.

The fact that touch responding by the older children was sensitive to outcome devaluation does not decide the issue because there is extensive evidence that a cue associated with an intrinsically attractive event also loses its attractiveness and hence its capacity to elicit approach responses when the associated event is devalued (e.g., Colwill & Motzkin, 1994). To determine whether a behavior is under instrumental control, one must show that it is mediated by the relationship between the action and its specific outcome rather than by an association between the outcome and a discrete stimulus, such as the butterfly icon. To this end, the relationships between the stimuli and outcomes need to be held constant, so only the specific action under investigation is uniquely related to the outcome that it produces. This condition is instantiated in the so-called bidirectional paradigm (Grindley, 1932), which is illustrated by the modification of our procedure used in Experiment 4.

During instrumental training, the children were presented with a single icon in the center of the screen that they first had to touch and then drag, either to the left or to the right, in order to produce video outcomes. Dragging the icon to the left produced video clips from one cartoon and dragging it to the right produced clips from the other cartoon. Although the initial touch response directed at the icon could have been elicited by the conditioned attractiveness of the icon through its association with the outcomes, the contingency between each drag response and its video outcome was entirely arbitrary. Consequently, the direction of the drag response could not have been mediated by a stimulus–outcome association and must have reflected a sensitivity to the instrumental contingency. At issue, therefore, is whether these actions were sensitive to outcome devaluation when tested in extinction. As only the older children showed an outcome devaluation effect in Experiment 2, we assessed the effect of the devaluation on the bidirectional drag response in children in the two older age bands.

Method

Participants

Thirty-two children (16 girls and 16 boys) aged between 27 and 48 months were divided into two 10-month age bands corresponding to those for Age Groups 2 and 3 in Experiment 2 (see Table 1). An additional 17 children in Age Group 2 and 13 children in Age Group 3 did not complete a full session and were replaced for the following reasons: interference by others or external events (8),
failure to acquire the required instrumental actions (6), failure to complete the minimum number of responses during instrumental training (6), equipment malfunction and experimenter error (5), abandoning the task because of distraction or fussing (5).

Procedure

Unless otherwise stated, the apparatus, stimuli, video outcomes, and general procedure were the same as those used in Experiment 2. In the present experiment, the instrumental actions required moving a small butterfly-shaped icon in different directions across the display. The display used during both training and test showed a small, 2 × 2-cm butterfly picture in the center of the screen, which served as a response manipulandum for the bidirectional action. When this square was touched, a smaller, red butterfly shape appeared under the child’s finger that could then be dragged across the touch-sensitive display as long as contact with the screen was maintained. If the butterfly icon was dragged far enough onto one of the two invisible trigger areas (each representing a vertical 5 × 16-cm strip) on the far left and far right of the display and then released, a video clip was presented. If the icon was dragged and released anywhere else on the screen, it disappeared and had to be picked up again from the center, and no video clip was shown. However, drags to the left and right that fell just short of the trigger areas and thus did not produce a video outcome were recorded as an instrumental response if they fell within the immediately adjacent area, which was of the same height as and half the width of each trigger area.

Action–outcome assignments and type of video devalued were counterbalanced across children in each age group but were always consistent for individual children so that, for example, a left drag would always produce clips from one cartoon, whereas a drag to the right would start a video clip from the other cartoon.

Pretraining. To train the children to perform the required instrumental actions, the experimenter demonstrated and explained the movement sequence to the child on the first two practice trials. Thereafter, the child completed three practice trials with each action. If necessary, the experimenter demonstrated the icon pick-up and icon drag sequence again during these trials. During pretraining, only two short, 8-s clips, one from each series, were used as outcomes.

Instrumental training. After pretraining, three new 12-s clips from each cartoon were used as instrumental outcomes. Initially, each response was trained separately. Half of the children were first trained to perform drags to the left, whereas the remaining half performed drags to the right first. In each case, one half of the screen was blacked out, so that only one of the two responses could be performed until five video outcomes had been obtained by dragging the icon in the relevant direction. After both responses had received this separate training, the whole screen was shown, and the child was free to make responses in either direction until at least four video outcomes from each series had been triggered or 4 min had elapsed, depending on which criterion was met first. In the latter case, to proceed to the next stage, a child had to produce at least two outcomes for each response to ensure that she or he had experienced both action–outcome contingencies. Including the five outcomes earned by each response during the single response training, this criterion ensured that the effect of outcome devaluation was assessed after each response had earned no less than seven outcomes but not more than nine.

Outcome devaluation. The outcome devaluation was similar to that used in Experiment 2. One set of three 12-s video clips was repeated five times with an interclip interval of 3 s. The devalued video outcome was counterbalanced with respect to the response to which it was assigned, left or right, and the cartoon from which it came.

Extinction test. Following outcome devaluation, the children were free to perform both actions for 2 min. However, neither response was followed by a video outcome. After the extinction test, the responses once again generated their respective outcomes for 2 min.

Results and Discussion

Instrumental performance was assessed during the period of instrumental training, in which the children were free to make both responses. As Table 2 shows, prior to devaluation there was no difference in the rate at which children performed the two actions. Following square-root transformation of the response rates, the F ratios for the effect of devaluation and the interaction of this variable with age were both less than 1. Age also did not significantly affect responding, F(1, 28) = 1.78, MSE = 0.17, p = .19. As Figure 3 and Table 3 illustrate, the children showed a devaluation effect in the extinction test by performing the response trained with the valued outcome relatively more than those trained with the devalued outcome, F(1, 28) = 7.86, MSE = 29.6. However, we can only be confident about the reliability of this devaluation effect for Age Group 3. Although there was no reliable interaction between age and devaluation (F < 1), planned comparisons revealed that the effect of devaluation was significant for the older children, F(1, 14) = 6.5, MSE = 27.1, but not for those in the younger Age Group 2, F(1, 28) = 2.14, MSE = 31.98, p = .17. The significant overall devaluation effect, which did not interact with age, is consistent with the claim that the children’s instrumental performance was primarily mediated by knowledge of the

![Figure 3. Mean percentage responding by the different age groups during the test for the responses trained with the valued and devalued outcomes. Error bars represent standard error.](image-url)
action–outcome contingencies. Even when both actions were directed toward a common response manipulandum, children over 3 years of age were able to encode a particular outcome in relation to a representation of the specific action that had produced it during training. Although the apparent devaluation effect in the test performance of children with a mean age of 31 months leads to the same conclusion for the younger children, the fact that the effect was not statistically reliable for Age Group 2 in the absence of an interaction with age makes the results difficult to interpret. If action–outcome learning did take place during training in both age groups, this learning was apparently less effective in controlling instrumental performance in the younger compared with the older group.

General Discussion

The main finding from this series of experiments is that children over 3 years of age are capable of goal-directed action as assessed by the outcome devaluation paradigm. In the absence of the outcomes during the extinction tests, these children performed both stimulus-directed responses (Experiment 2) and arbitrary bidirectional actions (Experiment 4) at a reduced level if the training outcome had been previously devalued. To do so, at least in the case of the bidirectional action, they must have encoded the specific contingency between action and outcome during training before integrating this information with a representation of the current value of the outcome. This capacity meets the criteria for goal-directed action offered by Dickinson and colleagues (Dickinson, 1985, 1989; Dickinson & Balleine, 1993).

There are two classes of psychological accounts of goal-directed behavior, the associative and the cognitive, both of which come in a variety of forms. Associative two-process theories assume that the pairing of the stimulus (S) with the outcome (O) establishes a stimulus–outcome association (S → O) through a Pavlovian learning process. The second, instrumental learning process then generates an outcome–response association (O → R) through experience with the response–outcome contingency, thereby allowing performance to be controlled by an S → O → R associative chain. On the basis of the assumption that the devaluation treatment generally decreases the excitability of the outcome representation, this model predicts reduced responding following devaluation. When expressed in terms of folk psychology, the sight of a butterfly icon made the children think of the associated outcome, which in turn made them think of the response that had produced that outcome. If outcome devaluation decreased the likelihood that the children were thinking of the devalued outcome, they would have been less likely to perform the devalued response. Whether or not two-process theory can explain the devaluation effects observed in the present experiments, and especially Experiment 4, depends on the nature of the instrumental learning process that generates the O → R link of the chain. According to Trapold and Overmier (1972), this link arises from the reinforcement of an association between the activated representation of the outcome that is excited by the stimulus and the response through the classic stimulus–response reinforcement mechanism (Hull, 1943; Thorndike, 1911). However, the demonstration of a devaluation effect with the bidirectional paradigm in Experiment 4 is problematic for this account. As the single visual icon was equally paired with both outcomes, each response would have been reinforced in the presence of activated representations of both outcomes so that devaluation of either of the outcomes should have had a comparable effect on both responses.

This problem does not arise in version of two-process theory espoused by Pavlov (1932) and his students (Asratyan, 1974; Gormezano & Tait, 1976), who assumed the O → R association arises directly from the pairing of each response with its outcome. As a consequence, each association is both outcome and response specific. A devaluation effect with the bidirectional response is also not problematic for the associative–cybernetic model of goal-directed behavior (Dickinson, 1994; Sutton & Barto, 1981; Thorndike, 1931). According to this theory, the sight of the butterfly icon causes the child to think of the two available responses, a left drag and a right drag, which would each in turn retrieve a thought of the associated outcome through the learned R → O associations. If the outcome is valuable, the theory assumes that the activation of the outcome representation feeds back on the response representation or thought to cause the child to perform it. Therefore, this theory assumes that a goal-directed action is mediated through an S → R → O associative chain, in which the activation of the response by the stimulus is not sufficient to generate overt responding without excitatory feedback from an activated and positively evaluated outcome representation. When expressed more colloquially, the R → O association enables the child to evaluate covertly the consequences of each action before deciding which one to perform.

Cognitive theories of goal-directed behavior also come in a variety of forms. There are a gamut of action theories (Greve, 2001), which assume that human goal-directed behavior is mediated by explicit, propositional-like representations of instrumental contingencies, whereas causal model theory argues that agents construct models of the causal efficacy of their actions (Waldmann & Walker, 2005). The scope of these cognitive theories is not restricted to human action and, in one form or another, they have also been applied to goal-directed behavior of nonhuman animals (Blaisdell, Sawa, Leising, & Waldmann, 2006; Dickinson, 1980; Heyes & Dickinson, 1990). It remains an important research issue whether the capacity of young children for goal-directed action depends on associative or cognitive processes.

In contrast to the oldest children, the sensitivity of the two younger age groups to outcome devaluation was less conclusive. Although the same overall pattern of results was obtained in the 27- to 36-month-olds, planned subgroup analyses in Experiment 4 failed to yield a significant devaluation effect in this group when a bidirectional response procedure was used. The rationale for using a bidirectional assay was to establish the mediation of responding by a representation of the action–outcome relationship rather than a stimulus–outcome association. Given that the devaluation effect was highly reliable for Age Group 2 in Experiment 2, which used a procedure that may have engaged stimulus–outcome learning, one possible interpretation is that the period between 2 and 3 years of age brings about a transition in behavioral control from stimulus–outcome learning to fully intentional goal-directed action.

For the youngest children, on the other hand, we could find no evidence that their responding was sensitive to outcome devaluation. Only when their responses actually produced the outcomes, as in the reacquisition test in Experiment 2, did they respond less for the devalued outcome than for the valued one, which estab-
lished that the devaluation treatment was effective for the 18- to 27-month-olds. Because the presentation of the outcomes during the reacquisition test may have reminded these children of the different values of the outcomes, we presented reminder cues for the video outcomes during the extinction test in Experiment 3 to parallel the conditions in the reacquisition test. However, the results indicated that the absence of a devaluation effect in the youngest group was not due to a failure of the outcome devaluation to transfer to the test of performance in extinction, as devalued and valued responses were performed at equivalent rates even when children’s responses to the reminder cues concurrently registered the differential value of the outcomes.

Although our procedure in Experiment 3 therefore showed effective transfer of current outcome value, it is possible that presenting the reminder cues throughout the instrumental test may not have been the most effective way to aid retrieval of children’s training memories. Effective reminder treatments have involved a pre-cuing procedure (e.g., Spear & Parsons, 1976), using either a salient cue from the original training context as in the present study (e.g., Rovee-Collier, Sullivan, Enright, Lucas, & Fagen, 1980), or a brief period of retraining administered in the original training context (e.g., Campbell & Jaynes, 1966; see also Adler, Wilk, & Rovee-Collier, 2000).

It is also possible that the youngest children’s insensitivity to outcome devaluation reflected the fact that they did not receive sufficient training to learn about the action–outcome contingencies. Although the duration of instrumental training was identical in all three age groups, the older children in Experiment 2 responded at a faster rate and therefore, on average, experienced a greater number of actions that were followed by an outcome (Age Group 2 M = 18.8, SE = 0.6; Age Group 3 M = 19.6, SE = 0.3) than the youngest children (Age Group 1 M = 16.9, SE = 0.9). Although the differences between these group means is not large, younger children may simply require relatively more training than older children for reliable action–outcome learning. As a result, the total number of responses made during training in the youngest group may still have failed to provide the critical level of exposure to the instrumental contingencies necessary for effective encoding of the relevant relationships at this age (see also Barr, Dowden & Hayne, 1996). In evaluating this explanation, however, it should be noted that animal studies of the effect of amount of training on sensitivity to outcome devaluation found that instrumental behavior is initially established as goal-directed and then becomes autonomous of the current goal value with further training (Adams, 1982; Dickinson, Balleine, Watt, Gonzalez, & Boakes, 1995; Holland, 2004; Killcross & Coutureau, 2003). Therefore, we consider it unlikely that the insensitivity of the youngest children to outcome devaluation was due to insufficient training.

Finally, it also seems implausible that the failure to detect a devaluation effect in the youngest children was due to a lack of power in our tests. A contrast between the performance of the devalued and valued actions of the children in Age Group 1 of Experiments 2 and 3 combined (n = 36) yielded a r value less than 1 with a power greater than .98 to detect a difference as large as that exhibited by Age Group 2 in Experiment 2.

In summary, the learning process that mediated responding by the youngest children in our task remains undetermined. Although the 18- to 27-month-olds appeared to have access to the current, relative values of the outcomes and clearly remembered how to perform the instrumental actions on test, their knowledge about the current incentive value of the outcomes did not seem to impact the processes responsible for the control of responding. As already noted, such resistance to outcome devaluation has often been observed in nonhuman animals following instrumental training even in the presence of the devalued outcome, as in Experiment 3 (Dickinson, Nicholas, & Adams, 1983), and especially following overtraining (e.g., Adams, 1982; Dickinson et al., 1995; Holland, 2004). On the basis of these findings, Dickinson and colleagues (Dickinson, 1985, 1989; Dickinson et al., 1995) have argued that instrumental training engages concurrently two different learning systems, one underlying goal-directed action and the other mediating stimulus–response learning, which does not encode the identity of the outcome. In human adults, the operation of this latter, habitual mechanism is clearly revealed when we make a slip-of-action because a particular stimulus elicits a well-trained but unintended response (Reason, 1992). Therefore, according to this theory, the goal-directed and habitual systems are in competition for the control of behavior, with the particular training history and test conditions favoring one system over the other.

The distinction between the goal-directed and habitual systems has recently been validated in animal studies by their dissociation in the rat prefrontal cortex (Balleine & Dickinson, 1998a; Corbit & Balleine, 2003; Coutureau & Killcross, 2003; Killcross & Coutureau, 2003; Ostlund & Balleine, 2005) and dorsal (Yin, Knowlton, & Balleine, 2004, 2005) and ventral striatum (Corbit, Muir, & Balleine, 2001). Critically, the acquisition of goal-directed action depends on the integrity of the prelimbic area of the rat prefrontal cortex (Ostlund & Balleine, 2005), which is thought to be homologous to dorsal prefrontal structures in the primate brain (Preuss, 1995; Rushworth, Walton, Kennerley, & Bannerman, 2004). Moreover, sensitivity to outcome devaluation is impaired by lesions of the orbital frontal cortex in monkeys (Baxter, Parker, Lindner, Izquierdo, & Murray, 2000; Izquierdo, Suda, & Murray, 2004).

This analysis raises the possibility that the actions of the youngest children were mediated by the habitual system, which accords with findings suggesting that 2-year-olds are significantly less skilled and more susceptible to perseverative, stimulus-controlled performance on problem solving and reasoning tasks than 3-year-old children (e.g., DeLoache & Sharon, 2005; O’Sullivan et al., 2001). Moreover, the development of the capacity for goal-directed action, at least as assessed by outcome devaluation, may depend on the maturation of the prefrontal cortex, which is known to undergo progressive development throughout early childhood (see e.g., Giedd et al., 1999; J. Huttenlocher, 1979; P. R. Huttenlocher, 1990, 2002; Jernigan et al., 1991; Mrzljak, Uylings, van Eden, & Judas, 1990; PfEffenerbaum et al., 1993; Sowell, Delis, Stiles, & Jernigan, 2001).

Nonetheless, the failure to find any evidence of truly goal-directed action in the youngest children in the present study would seem to conflict with claims made by numerous studies regarding the appreciation of goal-directed action in much younger infants and toddlers. On closer inspection, however, the conflict may be more apparent than real. First, while our studies concerned children’s appreciation of the linkage between their own actions and arbitrarily related, not currently perceived goal states, experimental studies of goal-directed action in infants have often investigated surprise reactions (assessed by looking time; for a review, see
Schöner & Thelen, 2006) to the anomalous (nonrational) acts of other agents (e.g., Gergely & Csibra, 2003; Woodward, 1998). For a recent attempt to explain discrepancies between abilities measured by looking time and the later capacity to take action or frame a judgment on the basis of such knowledge, see Russell (in press).

Secondly, demonstrations of insightful goal-directed behavior in older infants and toddlers have tended to employ the technique of deferred imitation, in which a novel action on a novel object is demonstrated on day one, not practiced at that time, and then evoked days or weeks afterwards (for a review, see e.g., Meltzoff, 2002). But although these studies tell us that children much younger than those who failed to demonstrate a devaluation effect here can remember and reproduce the particular form of another’s bodily act (Meltzoff, 1988) and may even appreciate why it was constrained to be that particular way (Gergely, Bekkering, & Király, 2002), no knowledge is being evoked of how their own actions, established by instrumental learning, relate to rewarding outcomes arbitrarily linked to them.

Moreover, when viewed in a different light, our data accord rather well with those from another area of cognitive–developmental research: the early development of executive function. Executive-inhibition tasks are typically those in which a prepotent lure has to be resisted while holding an arbitrary rule in working memory. It is generally claimed that between 3 and 4 years of age performance on executive-inhibition tasks becomes efficient (Perner & Lang, 1999; Russell, Mauthner, Sharpe, & Tidswell, 1991). However, when simpler executive-inhibition tasks are used, such as those requiring detour reaching plus an arbitrary act (Bı´ro, 2001; Russell, Judah, & Johnson, 2007), 3 years is the approximate age at which performance becomes less stimulus-controlled and less susceptible to the perseverative repetition of a habitual response.

Whatever the merits of the hypothesis that the transition in behavioral control to fully intentional goal-directed action occurs between 2 and 3 years, the present study has provided an empirical demonstration that by the time they are 3 years old, children are capable of true goal-directed action because they can pursue specific goals that are not in their immediate perceptual field in a way that is directly sensitive to changes in the current value of the goal and mediated by knowledge of the causal consequences of their actions. This capacity is an important component of becoming a fully autonomous intentional agent.

### References


Received September 11, 2006
Revision received June 5, 2007
Accepted June 12, 2007

---

**E-Mail Notification of Your Latest Issue Online!**

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at [http://notify.apa.org/](http://notify.apa.org/) and you will be notified by e-mail when issues of interest to you become available!