
http://researchrepository.murdoch.edu.au/20448
Task Structure Complexity and Goal Neglect in Typically Developing Children

Gareth Roberts and Mike Anderson

Neurocognitive Development Unit, School of Psychology

University of Western Australia

Word Count: 9093

Author Note

Gareth Roberts, Neurocognitive Development Unit, School of Psychology, University of Western Australia; Mike Anderson, Neurocognitive Development Unit, School of Psychology, University of Western Australia.

Gareth Roberts is now at the School of Psychology and Exercise Science, Murdoch University.

Mike Anderson is now at the School of Psychology and Exercise Science, Murdoch University.

The research was supported by an Australian Postgraduate Award from the Australian Government and conducted as part of the primary author’s doctoral dissertation. Special thanks to Kaitrin McNamara, Alex Nicol, Catherine Campbell, and Corinne Reid for assistance in participant recruitment and data collection.

Correspondence concerning this article should be addressed to Gareth Roberts, School of Psychology and Exercise Science, Murdoch, WA 6150. Email: g.roberts@murdoch.edu.au
Abstract

Goal neglect is a failure to enact task requirements despite being able to accurately report them. In this study, we introduce a new child-appropriate experimental paradigm to measure goal neglect in children between the ages of 7 and 11, and test the hypothesis that the complexity of an action plan, not real-time trial demands, increases goal neglect. Sixty-six children ($M_{\text{age}} = 9.50$) were administered a Feature Matching task. Half were given four rules for matching, and half three rules. After practice, the four-rule group was told to ignore the additional rule, and both groups completed an identical three-rule task. The results showed that the extra rule increased goal neglect and its correlation with fluid intelligence. While intermittent trial errors were correlated with fluid intelligence for both groups, only in the four-rule group were systematic rule failures (i.e., goal neglect) correlated with fluid intelligence. Task performance improved with chronological age, however when controlling for the influence of fluid intelligence, the relationship between age and task performance was effectively removed. This suggests a child’s current level of fluid intelligence (and not age) determines task performance. We suggest that the relationship between goal neglect, complex task instructions, and fluid intelligence is linked to the mental preparation for future events; i.e., mentally compiling verbal instructions into a set of activated goal representations in working memory that represent what is to be done and under what circumstances.

Keywords: Executive Function; Fluid intelligence; Goal Neglect; Task Complexity; Task Instructions
Task Structure Complexity and Goal Neglect in Typically Developing Children

Goals are "an intention to accomplish a task, achieve some specific state of the world or take some mental or physical action" (Altmann & Trafton, 2002, p. 39). Goals are central to the construct of executive functioning, which in turn is related to the construct of fluid intelligence (Friedman et al., 2006). Accordingly, many researchers directly or indirectly conceptualize the term executive function as a process of goal activation (Altmann & Trafton, 2002; Duncan, Emslie, Williams, Johnson, & Freer, 1996; Nieuwenhuis, Broerse, Nielen, & De Jong, 2004). In this context, executive functioning can be thought of as goal-directed processes that exert control over thought and behavior in novel and complex situations (Jurado & Rosselli, 2007). When goals are not brought to sufficient awareness, we can suffer from a lapse in intention, whereby nothing is attempted in behavior; despite verbal knowledge than an action is required (Duncan et al., 1996). This abulic dissociation, termed goal neglect, is similar to that seen in patients with damage to the frontal lobes.

Anecdotal and historical accounts of behavior after damage to the frontal lobes (Bianchi, 1922; Luria, 1966) suggest major defects in planning, coordinating or controlling a sequence of action. Behavior will often manifest as disorganized and fragmentary, with sequences of action left incomplete and purposeless actions introduced (Duncan, 1986). Intriguingly the same chaotic behavior can be seen in people with intact frontal functioning when they are faced with novel tasks of high complexity (Duncan et al., 2008). The best predictors of this type of behavior are tests of fluid intelligence, such as Raven’s Matrices or the Cattell Culture Fair test of $g$. These tests provide an excellent measure of Spearman's $g$, or general intelligence (Carroll, 1993). Spearman's $g$ derives from the finding that diverse tests of cognitive ability correlate with one another (Spearman, 1904). To explain the covariance in abilities, Spearman proposed that
some general factor ($g$) underlies individual differences in intelligence. Subject to vigorous study for over a century, $g$ is a well-established predictor of social, educational, neurocognitive and biological factors (Jensen, 1998).

Adult research regarding goal neglect has typically been conducted using two main experimental paradigms: the Letter Monitoring Task (Duncan et al., 1996) and the Feature Match Task (Duncan et al., 2008). These two experimental paradigms share several important characteristics that make them ideal for eliciting and measuring goal neglect. First, the tasks provide minimal performance feedback to participants. Once the experimental trials have started, participants are not provided with any information regarding their task performance by the experimenter or through the experimental paradigm. Second, the association between task stimuli and the corresponding action requirement is not explicitly stated by the task stimuli, nor is it obvious or intuitive in anyway. That is, the task requirements are ambiguous and hence participants must rely on internal representations to correctly guide behavior. Finally, the collection of task requirements and overall task structure must be novel. Novelty appears to be a key characteristic for both frontal patients and normal individuals in eliciting goal neglect (Duncan et al., 1996), with well-practiced habits and “crystallized” intelligence measures showing resilience against brain damage (Duncan, Burgess, & Emslie, 1995).

Children provide an excellent opportunity to investigate goal neglect due to the changes in executive function (Brydges, Reid, Fox, & Anderson, 2012) and fluid intelligence (Anderson, 1992) that occur throughout the childhood years. These changes are thought to be due to the development of the prefrontal cortex (PFC). Over childhood, the PFC undergoes substantial development (Casey, Giedd, & Thomas, 2000; Giedd & Rapoport, 2010; Giedd et al., 1999; Gogtay et al., 2004), reflected by marked changes in the abilities associated with the PFC
between the ages of 7 to 11 years (McArdle, Ferrer-Caja, Hamagami, & Woodcock, 2002). For example, between the ages of 7 and 11, children show an increased ability to hold information in mind and manipulate it (Diamond, 2002), which is a hallmark feature of goal-directed behavior and generally referred to as working memory (Baddeley, Sala, Robbins, & Baddeley, 1996). In a standard memory item of the Weschler Intelligence Scales, the forward digit span task, participants are asked to remember and recall a series of digits in the order in which they are heard. Children show an improvement of 1.5 digits at the task from ages 7 to 13. However, when children are required to recall the digits in the opposite order that they were presented, which requires manipulation of the temporarily stored information, there is an improvement of 3 additionally recollected digits (Diamond, 2002).

From this perspective, typically developing children can be considered frontally compromised compared to adults, and because of this they may provide important insights into goal-directed behavior that is dependent upon the PFC. As frequently pointed out by Luria (1966), young children often neglect task demands in situations that contain multiple rules, induce conflict between internal and external representations, or require the planning and execution of a multi-step behavioral sequence. A developmental paradigm frequently used to investigate these dissociations between knowledge and action is the Dimensional Card Change Sort task (DCCS) (Zelazo, Frye, & Rapus, 1996). In this paradigm, children aged between 3 and 5 are presented with bivalent stimulus cards and asked to sort the cards into piles based on the color or object identity of the stimulus on each card. While children initially perform appropriately by sorting the cards according to the experimenter’s instructions, when the instructions change and they are required to view the same stimuli in a new way, children characteristically behave inflexibly; a majority continues to act according to their original
responses rather than the experimenter’s new instructions. It is important to note that this effect is observed regardless of whether they were initially required to sort by color or shape. While the majority of 3-year old children fail to sort cards the post-switch cards correctly, by 5-years of age children show little difficulty in switching their sorting methods according to the change in task rules (Zelazo et al., 1996).

The DCCS has led to the development of theoretical models of cognitive development to explain this age-related change (Zelazo, 2004; Zelazo et al., 1996). The Cognitive Complexity and Control (CCC) model proposes that 3-year olds explicitly know both the pre-switch rules and the post-switch rules but without a higher order rule tying the two rule pairs together, they are unable to make a decision about which rule to use and thus persist in using the rule that is most strongly associated with the task (Zelazo et al., 1996). The CCC model posits that as children develop, they acquire the ability to embed the rule sets into a hierarchical structure. A theoretical extension by Zelazo (2004) argues that higher Levels of Consciousness (LOC) are essential in incorporating the two rule pairs into a single higher-order rule. Higher LOC are thought to develop with age and reflect increased re-entrant processing due to neural development of PFC. This increased “reflective consciousness” allows the contents of consciousness at one level to be considered in relation to other contents at the same level, which results in a more complex conscious experience and the ability to undertake more complex thought and action. Applying this model to age-related performance changes on the DCCS task, 3-year old children are thought to represent both the pre and post-switch rules at the same LOC. However, without reflection on the two rule pairs at a higher level of LOC, they are unable to integrate the rules and consequently show post-switch failures.
One of the first studies to specifically investigate goal neglect in children (Towse, Lewis, & Knowles, 2007) administered the DCCS, an inhibitory control task, and a child-appropriate version of the Letter Monitoring Task (Duncan et al., 1996) to a group of preschool children. To adjust the Letter Monitoring Task for use with young children, the rate of stimulus presentation was slowed down, the number of stimuli was reduced, and the task was made more motivating for young children through an engaging story and appropriate task stimuli (i.e., naming types of food to help a hungry bear, as opposed to arbitrarily naming letters). While these modifications made the task appropriate for testing young children, importantly they also preserved crucial features of the original adult task (i.e., minimal performance feedback, ambiguity, and novelty). Using the new task, Towse et al. (2007), investigated whether performance in the other two cognitive measures (DCCS and inhibitory task) predicted goal neglect to different rules of the Letter Monitoring Task. Using correlational analyses, the authors showed differential contributions from the two other cognitive measures on goal neglect on the child Letter Monitoring Task, but interestingly only the DCCS predicted performance on the component of the adult Letter Monitoring Task that is neglected by adult participants with low fluid intelligence. These results suggest the DCCS and child Letter Monitoring tasks share a central requirement of representing the goal state in an active state within working memory.

Within this goal maintenance perspective, Marcovitch, Boseovski, and Knapp (2007) tested the hypothesis that the ability to maintain goal representations in working memory may underlie performance on the DCCS. Drawing on adult research conducted with the Stroop task (Kane & Engle, 2003), the authors predicted that by frequently presenting “conflict” trials (trials requiring a switch between the stimulus dimensions) to children performing the DCCS, goal maintenance would be improved because children would be constantly reminded of the
appropriate goal state for the task. However, if “conflict” trials were infrequently presented, participants could slip into a state of responding where inappropriate actions would occur due to insufficient representation of the task goals in working memory. To test this prediction, the authors modified the DCCS by introducing a new “redundant” test card to the DCCS. These new cards were to be sorted the same regardless of the task rules, and were interspersed among the standard conflict test cards (cards that were sorted differently depending on the rules). The introduction of this new card type allowed the trial ratio of the DCCS to be manipulated. Using this new task, children aged between 4 and 5 demonstrated good performance when presented with a high proportion of “conflict” cards, consistent with the prediction that task context serves as a reminder of the relevant rules and keeps the goal state active. Furthermore, when children were presented with a high proportion of “redundant” cards, children showed poorer performance in sorting the “conflict” cards, an effect that persisted even when children were explicitly told the rule before each individual trial.

Using the same modified DCCS task, Marcovitch et al. (2010) tested whether working memory capacity (WMC) in children between 4 and 6 years of age predicted goal neglect. Using three working memory span tasks as a measure of WMC, the authors hypothesized that if the mostly redundant card sort condition relied on the maintenance of goal information in the face of interference, performance in the modified DCCS would be predicted by individual differences in WMC. This hypothesis was based on the theoretical prediction that goal neglect is a by-product of the inability to maintain goal representations over time and distraction, a skill directly measured by working memory span tasks (Kane & Engle, 2002). The authors further hypothesized that WMC would be related to performance on the mostly conflict sort for the younger children but not for the older children. This hypothesis was again, based on the
theoretical prediction that goal maintenance is still required for the mostly conflict card sort (Kane & Engle, 2003), but that older children generally do not require the additional scaffolding to succeed on the mostly conflict card sort, and thus no relationship between WMC and performance was expected. Both hypotheses were supported, WMC was a significant predictor of performance on both the mostly redundant card sort for both age groups, and the mostly conflict card sort showed a significant relationship with WMC for the 4-year olds but not the 6-year olds.

Although these developmental goal neglect studies emphasize the importance of maintaining the goal state, and how it can be affected through manipulations of task context, another important factor influencing goal neglect is the successful construction of a mental model from the task instructions presented to a participant. There is evidence from adult samples to suggest that the initial planning for future action from the task instructions imposes constraints on future successful performance in the task (Cohen-Kdoshay & Meiran, 2009; Duncan et al., 2008). For example, using a Feature Matching task, Duncan et al. (2008) presented half the participants with the standard rules to complete the task and the other participants with the standard rules plus an extra rule that related to a central stimulus feature that appeared on each individual trial. Even though participants were given practice with this task rule, they were not presented with any trials in the experimental sequence where it would actually apply. Both groups were given the same stimulus sequence and therefore an identical task. Despite performing precisely the same task, participants who were given the extra rule performed more poorly and found it difficult to satisfy the task demands. In addition, goal neglect in this condition was more strongly correlated with fluid intelligence, suggesting an important link between action planning and intelligence.
For this study we created a new paradigm to measure goal neglect in children aged between 7 and 11 years. The construction of a developmentally appropriate paradigm allowed for the investigation of factors that affect goal neglect in a sample already susceptible to action failures. We were specifically interested whether the adult findings of Duncan et al. (2008) could be replicated with a different task and in children. The new task was presented to children as a game requiring them to buy a house and a car for an alien named Yeebo. To help Yeebo buy a house and a car, they were told that when they see two pictures of a house or two pictures of a car to press the left or right key for the picture containing the lower price. They were warned, however, that an evil alien called Zorak was trying to mislead them and would sometimes present pairs of rockets. The children were informed that Yeebo did not need a rocket and did not want to waste his money, so they should withhold their response for these trials. The children were also told that Yeebo could only buy a house or a car, if the two objects were presented together. Half the children were also given an additional rule, that sometimes the two houses or two cars would be the same price, and for these trials they should press the down arrow key. This was the fourth rule of the experiment, and was not actually necessary for the experimental trials. After practice trials (that included this rule) these participants were told that they would not need this rule as all the cars and houses would be different prices. Children in both conditions were then presented with an identical stimulus sequence and hence an identical task.

In summary, we presented half the participants with a more complex body of task-relevant information but had all participants complete the same experimental sequence. As children are undergoing major neural changes in the PFC between the ages of 7 and 11 and show immature executive function (Tsujimoto, 2008), we sought to test Duncan et al.’s (2008)
instruction manipulation in a sample that showed high variability in fluid intelligence. It was predicted that the maintenance of a more complex task structure in working memory would lead to poorer task performance and a stronger relationship with fluid intelligence. Specifically, we hypothesized that participants who were given the extra rule (four task-rules) would show increased goal neglect compared with those who were only informed of three-task rules. It was also hypothesized that there would be a significant relationship between general task performance and fluid intelligence for participants in both conditions. However, it was hypothesized that goal neglect would be more strongly related to fluid intelligence in the four task-rules condition than in the three task-rules condition.

**Method**

**Participants**

There were 66 participants, aged between 7 and 11 years of age with \( M_{\text{age}} = 9.50 \text{ years} \) \((SD = 1.24)\) with an approximately equal gender balance (59.10% of the sample was male). Participants were typically developing and were from a medium to high socioeconomic background. Ninety-seven percent of participants were right handed. Participants were recruited through Project KIDS (Kids Intellectual Development Study). Project KIDS is an annual research program run by the Neurocognitive Development Unit at the University of Western Australia, Perth\(^1\). The program involves psychological testing embedded within an activity-day, holiday-program format that includes games and fun activities.

**Measures**

**Cattell Culture Fair Intelligence Test.** Cattell Culture Fair Scale 2A (Cattell, 1971; Institute for Personality and Ability Testing, 1973) was used as a measure of fluid intelligence.
The test assesses novel problem solving with geometric figures. There are four timed sub-tests (series completions, odd-one-out, matrices, and topology).

**Feature Match Task.** A child analogue of the feature match task was developed because pilot testing revealed the original feature match task from Duncan et al. (2008) was too difficult for children below the age of 12. The new task involved children being presented with two stimuli simultaneously in the center of the screen. The stimuli were of three objects (houses, cars and rockets), of two colors (red and blue) and each contained a dollar amount directly below them ($1, $2, or $3). The new experimental paradigm is shown in Figure 1.

Children were to respond if the two stimuli belonged to the same category, withhold their response if they belonged to a different category, and withhold their response if the stimuli belonged to the “No Go” category (rockets). If both stimuli were houses or cars, they responded by indicating whether the stimulus on the left or right contained the smallest dollar value via a button response (match rule). A response was only scored as correct if the correct dollar amount was selected. If the stimuli were a pair of rockets, they withheld their response (no go rule) and if the stimuli did not match (i.e., a house and a car), they withheld their response (no match rule). Half the children were given an extra rule that involved the dollar amounts below each object. They were told that if dollar amounts below both the stimuli were the same, they were to press another key (down arrow). The extra instruction was presented to children in the 4-rules condition as the first task requirement. Each task rule and its response requirements are outlined in Table 1. After completion of the Feature Match task, children were asked to repeat the rules of the task through a series of cued-recall questions (e.g., “What would you do if two rockets appeared on the screen?”). All children could successfully remember the task requirements.
Figure 1. Stimuli used in the child Feature Match task. The first trial is an example of the car match rule, the second trial is an example of the house match rule, the third trial is an example of the no-match rule and, the final trial is an example of the no-go rocket rule. For all rules, stimulus color and dollar amount were counter-balanced.
Table 1

**Rules for the child Feature Match task**

<table>
<thead>
<tr>
<th>Rule description</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Match rule If the two stimuli are both from the same category (i.e., two houses or two cars)</td>
<td>Press the left or right key for the stimulus with the lowest price</td>
</tr>
<tr>
<td>No Go rule If the two stimuli are rockets</td>
<td>Do not press anything</td>
</tr>
<tr>
<td>No Match rule If the two stimuli are not members of the same category (i.e. a house and a car)</td>
<td>Do not press anything</td>
</tr>
<tr>
<td>Unnecessary rule If the dollar amounts of the two stimuli are the same</td>
<td>Press the down arrow key</td>
</tr>
</tbody>
</table>

**Task Specifics.** The Feature Match experiment was programmed using functions of the Psychtoolbox running under MATLAB 2009b on a Windows XP PC. All stimuli were presented centrally and had an equal height and width of 3.5cm on a 19-inch CRT monitor (1280 x 760 resolution at 60 Hz). A trial constituted a pair of objects being presented for 1 s, followed by a 1 s blank interval. There were 4 blocks of 30 trials. Each block contained 18 no-match trials (stimuli that did not match), 9 match trials (stimuli that did match) and 3 no-go trials (stimuli that matched but required a response to be withheld). Directly below each object was a numeric value behind a dollar sign ($). These digits varied from one to three. All stimuli were either red or blue, and trials were equally both colors or mixed. Children were instructed to ignore color and focus on the objects. Responses were made via the “z” (left) and “/” (right) keys on the keyboard.
**Task Scoring.** For the purposes of analyzing the data we examined performance for each trial type in the task (match rule, no-match rule and no-go rule) and global composite scores, using two different scoring procedures: total error rate (reflecting general performance) and goal neglect (reflecting ability to adhere to task instructions). Total error rate was the total number of errors of any variety accumulated across blocks. Goal neglect was scored using a point system, where participants would obtain one point for each time a specific performance criterion was met for each task rule, and in each experimental block. This scoring procedure was taken verbatim from the Feature Match task of Duncan et al. (2008, experiment 4) and represents a neuropsychological scoring procedure. A rule was scored as “failed” if a participant made more than 25% incorrect responses for each rule in a task block. A participant’s total goal neglect score was their total score summed across all experimental blocks and for all task rules. In other words, this scoring procedure measured compliance to the task rules, i.e., whether a task requirement was being systematically ignored across the experiment. Therefore, a participant's total goal neglect score is conceptually different from a participant’s total error score. It is possible to make errors to different rules of the task but not meet the performance boundary to be counted as a rule fail. For example, errors could be made to each trial type of the task but not be frequent enough for any rule or task block to meet the criteria for goal neglect.

**Procedure**

In Project KIDS tasks completed by the children are incorporated into a theme where each task contributes to a group goal (Brydges et al., 2012). Children earn stickers or tokens that can be used to “buy” materials for the group goal. Each child only works on a task for a short period of time (25 to 30 minutes) before moving on to a different task, to ensure minimal fatigue and boredom. Children came to Project KIDS for two days of testing. All tests were
administered according to standardized procedure by trained staff. The experimental measures were administered in a group setting (six children) with one adult experimenter per child. Both the Cattell Culture Fair Intelligence test and Feature Match task were administered at different time periods over the two days for different groups of children.

Results

Data Screening

Visual inspection of the distribution of total errors revealed three clear outlier children who all had a total error rate of over 70%. These children were excluded from all analyses, as well as five other children who made multiple responses to single trials throughout the study. All the excluded children had recorded behavioral observations that they were not responding to the task appropriately, due to lack of motivation or fatigue. These exclusion criteria left a sample size of 58. Before scoring the data we examined group differences in the demographic information between children in the two experimental conditions. There were no significant differences for Cattell score, age, handedness or male-to-female ratio between the two samples (Table 2). Data from the Feature Match task was not normally distributed and hence not suitable for parametric statistics, and correspondingly nonparametric Mann-Whitney U tests were used for all between subject comparisons. One-tailed tests were used due to the strong a priori prediction that participants maintaining a more complex task structure would show poorer performance. For each test, exact p values and effect sizes (r) are reported.
Table 2

Demographic descriptives for both 3- and 4-rules conditions

<table>
<thead>
<tr>
<th></th>
<th>3-Rules (n = 28)</th>
<th>4-Rules (n = 30)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Cattell score</td>
<td>32.07</td>
<td>30.20</td>
</tr>
<tr>
<td>Age</td>
<td>9.65 years</td>
<td>9.36 years</td>
</tr>
<tr>
<td>Handedness</td>
<td>96% Right-handed</td>
<td>100% Right-handed</td>
</tr>
<tr>
<td>Sex</td>
<td>61% Male</td>
<td>57% Male</td>
</tr>
</tbody>
</table>

Task Structure Complexity and Individual Trial Type Performance

Table 3 shows the total number of errors made for each trial type between conditions and the results of Mann-Whitney U tests for each trial type. For all trial types, the maintenance of a more complex task structure was associated with poorer performance. Performance for each trial type across both conditions is shown in Figure 2. The instruction manipulation affected performance for all aspects of the task.

Table 3

Total error scores for each trial type for both 3 and 4-rules conditions

<table>
<thead>
<tr>
<th>Trial Type</th>
<th>3-Rules (M (SEM))</th>
<th>4-Rules (M (SEM))</th>
<th>z</th>
<th>p</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Match</td>
<td>6.04 (0.98)</td>
<td>10.73 (1.25)</td>
<td>-3.02</td>
<td>&lt;.01</td>
<td>.40</td>
</tr>
<tr>
<td>No Match</td>
<td>12.79 (3.62)</td>
<td>18.33 (3.11)</td>
<td>-2.12</td>
<td>.02</td>
<td>.28</td>
</tr>
<tr>
<td>No Go</td>
<td>1.21 (0.45)</td>
<td>2.17 (0.52)</td>
<td>-1.93</td>
<td>.03</td>
<td>.25</td>
</tr>
</tbody>
</table>
Task Structure Complexity and Global Measures of Task Performance

We then tested the hypothesis that children in the 4-rules condition would show poorer performance on average, than children in the 3-rules condition for our general measure of task performance (total error rate) and for goal neglect score. For total error score, the 4-rules condition ($M = 31.23$) had a $SEM$ of 3.78 while the 3-rules condition ($M = 20.04$) had a $SEM$ of 4.261. The results of the Mann-Whitney $U$ test were in the expected direction and significant, $z = -2.60, p = .004, r = .34$. For goal neglect score, the 4-rules condition ($M = 0.73$) had a $SEM$ of 0.20 and the 3-rules condition ($M = 0.21$) had a $SEM$ of 0.15, both of which were in the expected
direction and significant, $z = -2.50, p = .007, r = .32$. Both these differences constitute a medium
effect size and support the hypothesis that children in the 4-rules condition would show poorer
task performance compared to children in the 3-rules condition.

**Relationship between Task Performance, Fluid intelligence, and Number of Instructed Rules**

Having established that there was an experimental effect on the instructed rules
conditions, we now examined the relationship between task performance and fluid intelligence,
for both scoring methods and for each experimental condition. Non-parametric Spearman rank-
order correlations were performed between total error score, total goal neglect score, and fluid
intelligence, separately for both instruction conditions. For both the 4-rules and 3-rules
conditions, total number of errors was significantly correlated to fluid intelligence ($r_s = -.58, p =
.001$ and $r_s = -.69, p = <.001$, respectively). Interestingly, the only non-significant correlation
was between goal neglect score and fluid intelligence for the 3-rules condition, $r_s = -.18, p = .35$.
The corresponding correlation for the 4-rules condition was significant, $r_s = -.37, p = .043$.
Regression analyses between fluid intelligence and goal neglect in each condition confirmed this
difference; with the regression slope in the 4-rules condition being significantly different from
zero ($B = -.39, t (29) = -2.21, p = .04, R^2 = .15$) but not in the 3-rules condition ($B = -.075, t (27) =$
$- .38, p = .71, R^2 = .01$).

To explicitly test whether the relationship between fluid intelligence and goal neglect was
statistically significant between the two experimental conditions, we tested the interaction of the
two regression lines, under the null hypothesis that a single regression slope would be the best fit
for both experimental conditions. Despite the regression slope being significantly different from
zero in the 4-rules condition, and not in the 3-rules condition, the interaction was not significant,


$F(2, 54) = 2.41, p = .10$. This indicates that strength of the relationship between goal neglect and fluid intelligence was not statistically different between the two instruction conditions. As this non-significance may be due to a lack of statistical power, we used nonparametric bootstrapping analyses recommended for small sample sizes (see Preacher & Hayes, 2004; Preacher, Rucker, & Hayes, 2007) to further investigate the relationship between goal neglect and fluid intelligence between the instruction conditions. We computed 1000 bootstraps and calculated the 95% CIs using a bias-corrected and accelerated method (BCa) that adjusts for both bias and skewness in the bootstrap distribution (DiCiccio & Efron, 1996; Efron, 1987). For the 4-rules condition, the 95% BCa CIs between goal neglect and fluid intelligence were between $r_s = -.53$ and $r_s = -.20$, while for the 3-rules condition they were between $r_s = -.27$ and $r_s = -.13$. While the 95% CIs just overlap, the bootstrapping results suggest a consistent difference between the two experimental conditions in the strength of the correlation, with children in the 4-rules condition showing a stronger relationship between goal neglect and fluid intelligence.

**Relationship between Age, Fluid Intelligence and Task Performance**

As there is a well-known relationship between age and fluid intelligence during the childhood years, we investigated the effect of age on children’s total error score using sequential multiple regression. Specifically, we tested whether the relationship between age and task performance was due to the effect of fluid intelligence, rather than any broad developmental effects of chronological age. Entering age first into the model revealed that it was a significant predictor of total error rate, $F(1, 57) = 9.49, p = .003$, and explained 14.5% of the total variance ($R = .38$). However, entering fluid intelligence into the model in the second block of predictors revealed that fluid intelligence still contributed significantly to the prediction of total error score, ($B = -.64, t(57) = -4.33, p < .001$) and significantly improved model fit ($\Delta R^2$ change = .22, $F$
When fluid intelligence was entered first, it was a significant predictor of total error score $F(1, 57) = 31.57, p = .003$, and explained 36.1% of the total variance ($R = .60$) but when age was entered second, ($B = -.05, t(57) = -.36, p = .77$) it failed to predict any additional variance ($\Delta R^2 = <.01, F(1, 55) = .13, p = .72$). This result suggests that task performance is related to level of fluid intelligence, rather than to “general” processes associated with chronological age.

**Discussion**

The main aim of this study was to develop a child-appropriate Feature Match task for investigating goal neglect in children between the ages of 7 and 11. Secondary aims were to test whether the presence of extra instructions increases goal neglect, to test whether goal neglect is related to levels of fluid intelligence, and to test whether the relationship between goal neglect and fluid intelligence would be greater for children in the condition with the extra instruction. We wanted to test these hypotheses in children because of the well-known changes in executive function (Brydges et al., 2012) and fluid intelligence (Anderson, 1992) that occur between 7 and 11 years of age. It was hypothesized that children given four task rules as opposed to three would show poorer performance on the Feature Matching task, despite the extra task rule being unnecessary for successful performance. It was further expected that performance at the Feature Matching task would be correlated with fluid intelligence. The first hypothesis was supported; children given four rules showed decreased task performance (as measured by both total error rate and total goal neglect score). The second hypothesis was also supported; children in both instruction conditions showed a strong relationship between error rate and fluid intelligence. The third hypothesis was partly supported, while the explicit test of the hypothesis was non-significant, descriptively the correlation between goal neglect and fluid intelligence was higher
for children in the 4-rules condition than children in the 3-rules condition, and this was corroborated by separate regression and bootstrapping analyses.

Children who were given four task rules showed poorer performance on all rules of the task compared to children who were only instructed in the three necessary task rules. For both conditions, successful performance was significantly correlated with levels of fluid intelligence, reflecting the difficulty of complex goal-directed tasks that require the maintenance of multiple goals in working memory. These results are in accordance with research showing a strong positive correlation between WMC and fluid intelligence in children (de Abreu, Conway, & Gathercole, 2010; Cornoldi, Orsini, Cianci, & Giofrè, 2013; Giofrè, Mammarella, & Cornoldi, 2013; Hornung, Brunner, Reuter, & Martin, 2011) and in adults (Christal, 1990; Colom, Abad, Quiroga, Shih, & Flores-Mendoza, 2008; Fukuda, Vogel, Mayr, & Awh, 2010). This relationship is hypothesized to rely on the maintenance of “goal information” in the face of interference (Kane & Engle, 2002). The results of this study support this idea and demonstrate that the maintenance of a plan containing more task goals impedes performance. Moreover, the results suggest that the complexity of the plan held in working memory increases the strength of the relationship between goal failures and fluid intelligence. In addition, the results of the current study highlight the conceptual relationship between task complexity and fluid intelligence, which has been hypothesized and investigated for some time (Marshalek, Lohman, & Snow, 1983; Primi, 2002; Stankov, 2000) but has typically been confounded with complexities unrelated to the task structure. The current study utilized Duncan et al.'s (2008) instruction manipulation where the two factors can be separated. That is, the complexity of the task structure was manipulated purely through the verbal instructions presented to participants.
and not by increasing real-time demands of task execution (e.g., by increasing the span of an n-back task in a working memory paradigm).

Developmental analysis of the Feature Matching task revealed that task performance was negatively correlated with age, indicating that older children made fewer errors and showed less goal neglect. Interestingly, however, the performance improvement seen with age was not due to broad developmental effects associated with chronological age—when the influence of fluid intelligence (a non-age-standardized measure) was controlled for, the relationship between age and task performance was effectively removed. Therefore, although children improve at the Feature Matching task with age, this relationship was mediated by fluid intelligence. This could reflect individual differences among neurological variables that have complex non-linear relationships with age (Shaw et al., 2006; Shaw, 2007). In addition, this result suggests that goal neglect is a general phenomenon of task complexity, and that its behavioral manifestation will be similar across diverse age ranges, provided the participants are tested using an age-appropriate paradigm and have the same raw fluid intelligence test score. Support for this idea comes from Duncan et al. (1996) experiment 1, who used a broad age range of participants and found that the correlation between goal neglect and fluid intelligence was virtually unchanged when age was partialled out. Other goal neglect studies that have tested children (Marcovitch et al., 2007; Marcovitch et al., 2010; Towse et al., 2007), young adults (Roberts, Anderson, & Fox, 2013; Roberts, Jones, Ly, Davis, & Anderson, 2013) and elderly adults (Duncan et al., 1996) all indicate that goal neglect manifests similarly in different experimental paradigms.

Compared to the other goal neglect studies conducted on children (Marcovitch et al., 2007; Marcovitch et al., 2010; Towse et al., 2007) the current study differs in two main areas. First, the sample of children tested in this study was significantly older than all other studies, all
of which used children between the ages of 4 and 6. Indeed, considering past studies have
shown children beyond the age of 5 show little difficulty at the DCCS (Zelazo et al., 1996), the
Feature Match task introduced in the current study allows for researchers to investigate goal
neglect in an age appropriate manner for children between the ages of 7 and 11. Second, instead
of examining the relationship between WMC and goal neglect (e.g., Kane & Engle, 2003;
Marcovitch et al., 2010), we examined the relationship between fluid intelligence and goal
neglect. We argue that analyzing fluid intelligence is more informative and reflective of the
underlying theoretical framework behind goal neglect (Duncan et al., 1996; Duncan, 2010). In
particular, we refer to the fact that fluid intelligence tests consist of items that require novel
“offline” problem solving (i.e., no computer-controlled timing of stimuli, abstract stimulus-
response mappings, etc.). Despite this difference, fluid intelligence tests correlate very highly
with experimental paradigms that test the ability to plan for future action “online” (e.g., goal
neglect paradigms). In contrast, WMC paradigms do not contain any overt problem solving, and
are more akin to “online” goal neglect paradigms where performance is due to the real-time
control of action in a highly controlled environment.

In the current study, task performance was scored using two different methods and the
difference between the methods is conceptually important. Traditional goal neglect (as scored in
Duncan et al., 2008) can be thought of as the competency to implement the required task model
and reflect the participant’s understanding of the task’s requirements. On the other hand,
standard rule errors are tantamount to errors that reflect momentary “mistakes” or “slips” made
by the participant. Using this scoring procedure meant that it was possible for children to show
no goal neglect—that is, the child was capable of maintaining the task model and generally
complying with task requirements (they understood and could “do” the task) but still make
errors. The results of the current study show that this variability of momentary mistakes and errors across the task, while distinct from goal neglect, is highly related to fluid intelligence and from any theoretical perspective this would be an unsurprising result. These errors simply confirm the difficulty of novel multi-rule experimental paradigms. Interestingly, however, goal neglect was only related to fluid intelligence in the extra instruction load condition. While the explicit test of this hypothesis was not statistically significant; bootstrapping analyses revealed a consistent difference in the strength of the correlation, and separate linear regressions showed that the slope was significantly different to zero in the 4-rules condition but not in the 3-rules condition. These results in combination with the small-to-medium effect sizes found in the between-subjects comparisons of task performance, suggest that a future better-powered study would be more sensitive in detecting this effect, and help in elucidating on the distinction between momentary errors versus systematic rule failures.

The performance decrement observed by participants in the 4-rules condition is similar to a well-known effect in the prospective memory literature. That is, the cost of a secondary task on a primary on-going task even if the second task is never actually encountered. In a seminal paper, Smith (2003) had half of the participants perform a lexical decision task with an embedded prospective memory task, and half perform just the lexical decision task. Upon analysing response times, the author found that participants who completed the embedded prospective memory task had longer response times on non-prospective memory target trials than participants performing the lexical decision task alone. These results suggest that the maintenance of an action intention requires capacity either for the storage of the intention or for the checking of individual stimuli. Another related field of literature concerns the implementation of verbal instructions into S-R mappings. For example, there is evidence that
verbal instructions are rapidly converted into mental representations that can immediately influence cognitive control, somewhat like a “prepared reflex” (Cohen-Kdoshay & Meiran, 2007; 2009). For example, Wenke, Gaschler and Nattkemper (2007) found that when participants were instructed of S-R mappings for a size task, and then told that they would be doing a location task; participants suffered interference effects when the stimuli were incongruent with the first task but congruent with the second. This result suggests that participants had formed a mental representation of the first task that was subsequently affecting their behavior. Similarly, participants in the current study appear to have rapidly converted the verbal instructions into active goals that interfered with successful task performance.

The results of the current study suggest that in a complex, goal-directed situation the cognitive system attempts to guide behavior by any representation it has available. This reinforces the fallibility of goal representations (Davelaar, 2011) and is consistent with theoretical computational models of cognitive development that emphasize active and latent representations guiding behavior, with higher-order representations serving as a goals that bias lower-level processes toward the goal at hand (e.g., the computational model by Morton & Munaka, 2002). Further computational modelling may shed light on the dissociation of declarative and procedural knowledge. For example, how the cognitive system can reproduce the verbal rule instruction but not be able to keep these representations at a high enough level of activation to guide behavior. Indeed, the model by Morton & Munaka (2002) suggests that representations are not all-or-none but rather vary continuously in their strength. This “graded representations” account suggests dissociations between knowledge and action result because weak prefrontal representations suffice for some tasks but not others. Further research is needed
to formulate explicit mechanistic accounts of how extra task representations negatively affect performance.

The results of the current study emphasize the advantage of looking at goal-directed behavior as a holistic picture, and not just at isolated aspects of cognitive control (Duncan, 2010). The finding that an extra task instruction can lead to detrimental performance for typically developing children in all trial types for a difficult goal-directed problem reinforces this perspective. Each trial type in the Feature Match task consisted of its own difficulties and processes, and indeed, one could argue that the complete task requires commonly posited executive functions such as task shifting, inhibition of prepotent responses, and the updating of working memory (Miyake et al., 2000). However, we argue that by focusing on isolated elements of control; the larger picture of successful goal-directed behavior is clouded. Moreover, recent structural equation modeling work in children has shown that a unitary executive function factor provides the best fit to the data in typically developing children between the ages of 7 and 11 (Brydges et al., 2012). That is not to say that elements of “cognitive control” cannot be meaningfully subdivided into components, but rather that the orchestration of a complex multi-step sequence of behavior appears to be most closely aligned to what is measured by tests of fluid intelligence (Duncan, 2010) and that this may be a key component to their success in predicting real world outcomes (Gottfredson, 1997).

In conclusion, the current study showed that goal neglect is highly related to fluid intelligence in a sample of typically developing children between the ages of 7 and 11. To do this we introduced a new developmentally appropriate paradigm for investigating goal neglect that required children to match stimuli using different rules. The results highlight the strong relationship between the construction of plans for future action and cognitive processes measured
by tests of fluid intelligence. The results were obtained by instructing half the children of four rules for a matching task, and the other half with three rules. For the experimental trials, only three rules were ever used, hence children were presented with an identical stimulus sequence. Therefore, the only difference between the two conditions was the action plan that was created to prepare for the task. This had a differential effect on goal neglect and the probability of making errors or mistakes. The latter was related to fluid intelligence in both conditions, but goal neglect itself was only related to fluid intelligence in the condition where an individual’s mental representation of the task embodied an extra (but unnecessary) task instruction. We suggest the extra task instruction increases goal neglect due to a disruption in the ability to convert all task rules into a set of sufficiently activated goal representations that can guide behavior.
Footnotes

1 Now relocated to School of Psychology and Exercise Science, Murdoch University, Western Australia.
References


