**Documentation for the data file McCormack McGourty Simms Beckers JECP**

The data file McCormack McGourty Simms Beckers JECP reports data from an experiment published as:

McCormack, T., Simms, V., McGourty, J., & Beckers, T. (2013). Blocking in children’s causal learning depends on working memory and reasoning abilities. Journal of Experimental Child Psychology, 115, 562–569.

Full details of the experimental methodology are provided in the published paper; a key to each column of the data file is provided in the file.

**Experiment**

Ninety nine children completed a causal learning task that was an analogue of the food allergy paradigm used with adults. The cue competition effects of blocking and unovershadowing were assessed under forward and backward presentation conditions. Children also answered questions probing their ability to make the modus tollens inference posited to be necessary for blocking by a reasoning account of cue competition. For the first time, children’s working memory and general verbal ability were also measured alongside their causal learning. The magnitude of blocking and unovershadowing effects increased with age. However, analyses showed that the best predictor of both blocking and unovershadowing effects was children’s performance on the reasoning questions. The magnitude of the blocking effect was additionally predicted by children’s working memory abilities. These findings provide new evidence that cue competition effects such as blocking are underpinned by effortful reasoning processes.

**Participants.** Ninety nine children were recruited for the study, 53 boys and 46 girls (*M* = 65.57 months, Range = 48-83 months). Fifty children took part in the forward order condition and 49 in the backward order condition, with children randomly assigned to conditions.

**Materials.** A toy robot purposely built for the causal learning task was used. Toy foods could be fed to the robot by placing them on a platform in his mouth; pressing the robot’s nose caused the platform to move downwards and tilt, dropping the foodstuffs into the robot’s ‘tummy.’ On releasing the robot’s nose the platform moved upwards into its initial position inside the mouth. Foodstuffs placed into the robot’s tummy via the platform were visible as the tummy area was covered by a transparent Perspex center. Along either side of this area were semi-opaque light boxes containing battery-powered LED lights. Each light box on either side of the tummy was divided into two, with the bottom half of each light box colored pink and the upper half red. A concealed speaker was also contained within the robot. When the platform moved downwards and tilted to place a food into the tummy, either a weak, strong, or no response occurred. A weak response consisted of the bottom half of the tummy lighting up along with a low, quiet sound. The strong response consisted of all of the tummy lighting up (bottom and upper halves) accompanied by a higher, louder sound. The duration of these responses was 3 s, following which the experimenter retrieved the foodstuffs from a hole at the back of the robot. All responses were predetermined by an input file selected by the experimenter from a computerized control program on a laptop. Seven sets of five toy foods (e.g., bread, an apple, yoghurt) were used, with one set used as a training set.

Working memory was measured using a computerized version of the animal recall task similar to that used by Barrouillet, Gavens, Vergauwe, Gaillard, & Camos (2009); the stimuli were presented on a laptop using E-Prime 2.0 software. In this task children were shown a series of pictures of animals and asked to recall the animals in the correct order*.* Children were required to name the color of smiley faces that appeared on the screen in between the pictures of the animals. Four levels of increasing difficulty were created, with the first level requiring one animal to be recalled while the fourth level involved recalling four animals. There were four trials in each of the four levels. There was no repetition of animal pictures to prevent between-trial interference; hence 46 pictures of different animals were used across training and test trials. In each trial a smiley face coloured red, blue, or yellow was also presented between the animal pictures. Each animal picture and smiley face was approximately 6.5 cm X 6.5 cm. Animal pictures appeared on the screen for 2000 ms, followed by a blank screen for 500 ms, followed by a smiley face for 2667 ms. A further blank screen appeared for 1333 ms before the next animal picture appeared. After each trial a question mark appeared in the centre of the screen prompting the child to recall the animal/animals shown. Once the child had responded, the experimenter would initiate the next trial, with the word ‘ready?’ appearing on the screen before the first animal.

A modified version of the digit recall task from the Weschler Intelligence Scale for Children IV (WISC IV, Weschler, 2003) was also used. The British Picture Vocabulary Scale II (BPVS II,Dunn, Dunn, Whetton & Burley, 1997) was administered as a control measure of verbal ability. In this standardized task, children are shown sets of four pictures and, for each set, have to choose a picture that matches a word spoken by the experimenter. This task is extensively used in the UK as a measure of receptive vocabulary.

**Procedure**

Children who were tested in their schools took part in two separate testing sessions that were conducted on the same day, each lasting approximately 25 mins. The first session involved the first half of the robot task followed by the animal recall task. The second session included the second half of the robot task, the digit span task and the BPVS II. Children who were tested in the laboratory completed the tasks in a single testing session lasting 1-1 ½ hours, but breaks were incorporated in the session between tasks.

**Robot task.**

***Pretraining.*** Initially participants were shown how the robot ‘eats’ food which then goes into his ‘tummy.’ Pretraining followed the procedure outlined in Table 1a. Demonstrations were given in a fixed order illustrated in the table, with each demonstration being repeated. The foods used in the pretraining phase remained constant for each participant. The occurrence of a weak outcome was described as “a bit of the robot’s tummy lights up” whereas a strong outcome was referred to as “all of his tummy lights up.” The pretraining phase included a number of questions to ensure that the child understood the nature of the task.

***Training.*** Training trials used during testing are shown in Table 1b, with each participant completing 6 tasks. Participants were randomly assigned to either a forward or backward group; forward group participants completed three forward blocking and three forward unovershadowing tasks while those in the backward group completed three backward blocking and three backward unovershadowing tasks. The presentation order of blocking and unovershadowing tasks was counterbalanced across participants. As in previous studies with children, the cue competition effects of blocking and unovershadowing were assessed in separate tasks; this was to ensure that children did not have to sit through a very large number of trials per task. There were two training phases to each task during which either elements (single food cues) or compounds (two foods together) were shown with their associated outcome. During the training phase, each trial was shown three times and the order of presentation of trials within each phase was varied. So, for example, in a forward blocking task, participants would see a trial involving the robot being fed one foodstuff that made the robot’s tummy light up (A+) and another trial in which a different foodstuff did not make the robot’s tummy light up (E-) in Phase 1. Participants saw each of these trials three times in a randomized order. In Phase 2, trials consisted of participants being presented with the compound cues (AB+ and CD+) three times in a pseudo-randomized order. Thus, for each task participants observed 12 training trials in total. Different novel sets of food stuffs were used for each separate task. Foodstuffs in any given set were counterbalanced in terms of which element they represented (A-D). However, in any set the foodstuff representing element E remained constant; this was a filler item that was included to ensure that there was at least one cue per task that was not paired with an outcome.

***Test phase.*** Following training, children were asked test questions in a counterbalanced order: “Is (food name B, e.g., an apple) a food that makes the robot’s tummy light up?” and “Is (food name C, e.g., bread) a food that makes the robot’s tummy light up?” Children were required to give a yes or no answer to each of these questions. Children were also asked the following forced-choice question: “If you had to choose one of these foods to make the robot’s tummy light up, which one would you choose (experimenter holds out B and C)?” Children indicated their choice by either pointing to or naming one of the foods. Feedback was not provided to participants. All six tasks involved an identical training and test process. Children received the first three tasks in one testing session, and at the start of the second testing session they were provided with a brief reminder of the training that they had experienced before completing the last three tasks.

***Reasoning questions.*** In addition to assessing cue competition effects, children’s ability to make the relevant modus tollens and disjunctive inferences was also assessed. At the end of each testing session, children answered one modus tollens question and one disjunctive reasoning question (i.e., they received each question twice, one in each session, but with new foods for each task). For the modus tollens question, children were initially shown a pair of foods fed to the robot a single time, and part of the robot’s tummy lit up. Children were then asked “Does only one of these foods make the robot’s tummy light up, or do both of these foods make the robot’s tummy light up?” The order in which “only one” or “both” of the foods were mentioned was counterbalanced. For the disjunctive reasoning question, children were again shown a pair of foods fed to the robot a single time, with part of the robot’s tummy lighting up. The experimenter then said “One of these foods isn’t a food that makes the robot’s tummy light up. Is the other food one that makes the robot’s tummy light up?” Children received a score varying from 0-2 on the modus tollens question and from 0-2 on the disjunctive questions.

**Animal Recall.** At the start of the training phase, children saw six colored smiley faces and were asked to name the colors. Six animals were then shown with the child being asked to repeat the name of the animal after the experimenter. Following this initial training there were two practice phases: two 1-animal trials and two 2-animal trials. Each trial in the first practice phase proceeded as follows: one animal appeared on the screen, the experimenter named the animal and the child repeated the name after the experimenter, then the child named the color of the smiley face that appeared next, following which they were prompted to recall the animal that they had seen. Children completed two practice trials of this sort; however if they incorrectly recalled the animal on either trial the experimenter reminded the child of which animal they had seen and repeated the trial. The second phase of training with 2-animal practice trials was then given once the 1-animal practice trials were correctly recalled. The necessity to recall the animals in the correct order was introduced in this practice phase, with the experimenter asking the child to recall “what animal you saw first and what animal you saw next” in the first 2-animal practice trial. The need to recall animals in order was reinforced following the child’s response. If the child recalled correctly, the experimenter said“Great, that’s right, we did see the X first and then the Y”. However if the child recalled the animals in the wrong order the experimenter said“Almost! But remember we saw the X first and *then* the Y”.If incorrect responses were given to either of the two trials in the second practice phase, this phase was repeated to ensure the child understood the task.

Testing immediately followed the practice trials. The first level of test trials consisted of four 1-animal trials and followed the same procedure used in the practice phases, with subsequent levels involving four 2-animal, 3-animal and 4-animal trials. If all of the animals in one out of the four trials in a level were recalled, even in the incorrect order, the child moved onto the next level with trials of increased length. The task was terminated when a child failed to recall all of the animals in any of the four trials at a particular level. Children received a score of .25 for each animal recalled in the correct list position. Thus, for the first level of 1-animal trials, children could score a maximum of 1 (4 lists x 1 item x .25), and for the next level children could score a maximum of 2 (4 lists x 2 items x .25).

**Digit Recall.** A modified version of the WISC-IV digit span sub-task was used, with the number sequences presented to each child remaining constant. The digit sequences ranged from 2-9 digits in length with eight difficulty levels and one practice level. Each level consisted of a particular digit span length being presented for two trials with a third trial available but not automatically administered. The task was terminated once a child got both trials in a particular level incorrect. If a child got one trial of a particular length correct and the other incorrect in the same level, the third trial was given. If this third trial was recalled correctly the task continued and the next trial length was given. However, if the third series was incorrectly answered the task was terminated as the child would have incorrectly answered twice on the same digit length. Following the WISC manual, a score of 1 was given for each correct trial recalled in the correct order, so the maximum score for an individual level of a particular length was 2.

**Results.**

The data file contains the data for all measures reported in the paper as follows:

* Age in months
* Raw scores on the British Picture Vocabulary Scales
* Standardized scores on the British Picture Vocabulary Scales
* Digit span (memory measure)
* Animal recall scores (memory measure)
* The average Z scores for the two memory measures
* Blocking difference scores, which indicate the extent to which children showed blocking in the causal learning task.
* Unovershadowing difference scores, which indicate the extent to which children showed unovershadowing in the causal learning task
* Scores on the modus tollens reasoning questions (0-2)
* Scores on the disjunctive reasoning questions (0-2)