

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20

Dissecting children’s observational learning of complex actions through  
selective video displays

Running head: Dissecting children’s observational learning

Date of submission: 21/09/2012      Date of Resubmission: 03/06/2013

Word count of main text excluding abstract, references, tables and figures: 8178

21 **Abstract**

22 Children can learn how to use complex objects by watching others, yet the relative  
23 importance of different elements they may observe, such as the interactions of the  
24 individual parts of the apparatus, a model's movements, or desirable outcomes,  
25 remains unclear. One hundred and forty 3-year-olds and one hundred and forty 5-  
26 year-olds participated in a study in which they observed a video showing tools being  
27 used to extract a reward item from a complex puzzle box. Conditions varied according  
28 to the elements that could be seen in the video: (i) the whole display including the  
29 model's hands, the tools and the box, (ii) the tools and the box but never the model's  
30 hands, (iii) the model's hands and the tools but not the box, (iv) only the end-state  
31 with the box opened, and (v) no demonstration. Children's later attempts at the task  
32 were coded to establish whether they imitated the hierarchically-organised sequence  
33 of the model's actions, the action details and/or the outcome. Children's successful  
34 retrieval of the reward from the box, and the replication of hierarchical-sequential  
35 information were reduced in all but the "whole display" condition (i). Only once a  
36 child had attempted the task, and witnessed a second demonstration, did the display  
37 focused on the tools and box (ii) prove to be better for hierarchical-sequence  
38 information than displays that focused on the tools and hands only (iii).

39

40 Dissecting children's observational learning of complex actions through  
41 selective video displays

42 Observational learning allows a child to acquire much adaptive information  
43 from his or her cultural environment, and several different processes of learning  
44 underpin the assimilation of the critical aspects of what is witnessed. Tomasello,  
45 Kruger and Ratner (1993) distinguished different forms of observational learning  
46 including mimicry, in which the actions of another individual are copied with little  
47 thought to the resulting outcome, and imitation, where an individual instead  
48 reproduces the outcome, as well as the actions that lead to the outcome. Whiten and  
49 Ham (1992, page 250) defined imitation more simply as a process in which "B learns  
50 some aspect(s) of the intrinsic form of an act from A". By contrast in emulation an  
51 observer focuses on the mechanics of a scene, potentially learning about the  
52 affordances of the objects concerned (Byrne, 1998), for example that an object can be  
53 moved in a certain manner (object movement re-enactment, Custance, Whiten, &  
54 Fredman, 1999), or that a certain goal can be achieved (goal emulation, Whiten &  
55 Ham, 1992).

56

57 *Dissecting imitation versus emulation*

58 There has been a recent drive in both comparative (Whiten, Horner, Litchfield,  
59 & Marshall-Pescini, 2004) and developmental psychology (Want & Harris, 2002;  
60 Whiten, McGuigan, Marshall-Pescini and Hopper, 2009) to dissect these different  
61 mechanisms within the observational learning process to establish the importance of  
62 each (see Hopper 2010 for a review). To do this, a number of ingenious paradigms  
63 have been developed, two of which are particularly relevant to the current study: two-  
64 action tasks and ghost controls.

65           In two-action tasks (Dawson & Foss, 1965) the same outcome is achieved by a  
66 model or models using either of two alternative methods, such as pushing a lever  
67 versus pulling a lever. Replication of the method a participant saw a model use to  
68 achieve the outcome then implies imitation. Achieving the outcome witnessed, but not  
69 using the method observed, implies result or goal emulation. Research using such  
70 two-action tasks has been extremely fruitful, showing that young children often  
71 imitate, copying the means they see others use to achieve a desirable outcome  
72 (McGuigan, Whiten, Flynn, & Horner, 2007; Nielsen & Tomaselli, 2010; Tennie,  
73 Call, & Tomasello, 2006). However, in some contexts children have also been shown  
74 to be selective learners, either not copying all the actions they have witnessed or  
75 replicating the outcome but using alternative means (Bekkering, Wohlschläger, &  
76 Gattis, 2000; Carpenter, Akhtar, & Tomasello, 1998; Flynn, 2008; Meltzoff, 1995;  
77 Wood, Kendal, & Flynn, 2012).

78           “Ghost control” experiments instead remove the agent from the display  
79 witnessed, so that it takes a “ghostly” form, offering participants the opportunity to  
80 recreate the outcome they witnessed through emulation, as the pertinent parts of the  
81 apparatus move but the absence of an agent offers no possibility of imitation. Such  
82 displays have been engineered through use of a remote control (Thompson & Russell,  
83 2004), by the discreet use of fishing line (Hopper, Lambeth, Schapiro, & Whiten,  
84 2008; Hopper, Flynn, Wood, & Whiten, 2010; Tennie *et al.*, 2006) or by digitally  
85 altering a video display (Huang & Charman, 2005). Children as young as 17 months  
86 have been found to learn from displays that present only information about the  
87 interactions of objects, without the model’s movements being seen (Huang &  
88 Charman, 2005). By contrast, Tennie *et al.* (2006) found that 18-month-olds did not  
89 match the pushing or pulling of a door when it was displayed within a ghost control,

90 whereas they did match the method witnessed when the model was included in the  
91 scene. Twenty-four month-olds matched the method witnessed in both conditions.  
92 Finally, Hopper and colleagues tested slightly older children's (three- and four-year-  
93 olds) learning with (i) a simple bi-directional task in which a door could be moved to  
94 the left or right (Hopper *et al.*, 2008), and (ii) a more complex tool use task, in which  
95 a tool could be used to remove an obstructing block, both resulting in the release of a  
96 reward (Hopper *et al.*, 2010). For the bi-directional door task children matched the  
97 response they witnessed on the first trial. But when all responses were considered,  
98 only children who witnessed a whole demonstration or an "enhanced" ghost control  
99 (in which another child was present but did not manipulate the apparatus) matched the  
100 direction that they witnessed the door moved above chance levels. The enhanced  
101 control provided a social element to draw the participant's attention to the display,  
102 working under social facilitation to control for mere presence effects (Akins, Klein, &  
103 Zentall, 2002; Fawcett, Skinner, & Goldsmith, 2002; Klein & Zentall, 2003). Children  
104 who witnessed a standard ghost control, with no other child present, did not match the  
105 direction of the door movement. The tool use task produced similar results, with  
106 children who witnessed a whole demonstration showing the best performance, and  
107 children in a ghost control showing better performance than children who witnessed  
108 no demonstration.

109         In the present study we developed a new approach to dissecting such elements  
110 by creating videotapes that revealed different aspects of the execution of a complex  
111 task. We compared the level of success on this task and the specific components of a  
112 witnessed demonstration that were copied when children were presented with  
113 information focussed only on the affordances or movement of the apparatus  
114 (addressing the role of information that underpins emulative learning), or on the end-

115 state of the task (addressing the role of information that underpins goal emulation), or  
116 on the physical actions made by a model (addressing the role of information that  
117 underpins the imitation of body movements). During observational learning observers  
118 may process one or several aspects within such a scene. By presenting displays that  
119 isolate different aspects within a scene, at the detriment of access to other forms of  
120 information, we can establish how these drive different social learning processes, and  
121 as a result can establish how important each aspect is to a child's success and his/her  
122 adoption of aspects of the demonstration. For example, in a ghost control  
123 demonstration children are presented with all the functional information with regard  
124 to an object's affordances and the inter-relations between objects and, objects and  
125 tools. In an end-state condition, an observer can see the result that can be achieved,  
126 but must infer how to achieve this end-state through their own endeavour. In one of  
127 our conditions children witness the hands of a model using a series of tools but never  
128 see the tools connect with the main apparatus, thus presenting body movements and  
129 tool manipulations (but not allowing access to information about the tool to object  
130 connection); thus assessing how much children's success is through observation and  
131 replication of the bodily actions. Each of these conditions is compared to a whole  
132 display in which an observing child sees the model, the tools and the task during the  
133 demonstration. Comparisons with this whole display allow one to examine the driving  
134 power of each of these elements (object manipulations, end-state, body movements) in  
135 a child's observational learning. For example, if a similar level of successful retrieval  
136 of a reward or fidelity to specific elements within the display is achieved in one of the  
137 manipulated displays compared to the whole display, then we can be sure that this  
138 form of information, and the learning process it allows (tools and box only, emulation;  
139 tools and hands only, bodily imitation; end-state, goal emulation), is a significant

140 driver in observational learning. Further, by comparing the different manipulated  
141 displays, we can establish whether one is more influential in children’s observational  
142 learning than another.

143

#### 144 *The Task*

145 Many observational learning studies have been concerned with establishing  
146 *how early* children begin to replicate observed information, and so have used actions  
147 applied to rather simple tasks, like placing a block into a hole or opening a box by  
148 sliding a door left or right. While an analysis of observational learning in these  
149 contexts can tell us much about such processes in infants, it is also essential to  
150 understand how observational learning occurs as children develop and experience  
151 more complex action sequences. Thus the current study extended previous research by  
152 addressing how different forms of information affect a child’s ability to learn to  
153 complete a complex task. The task used was an “artificial fruit” (Whiten, 1998),  
154 designed as an analogue of a tool-based naturalistic food processing task faced by  
155 apes and children alike. Our “Keyway Fruit” (KW, see Figure 1) was a puzzle box  
156 requiring the execution of sixteen consecutive actions to extract a reward held inside.  
157 The new and critical aspect of our study was that we showed children video displays  
158 that differed in terms of the type of information available. Some children witnessed  
159 the whole display with the KW, tools and hands of the model manipulating the tools  
160 (thus providing a benchmark from which to comparing to other conditions), in a  
161 second condition children saw only the KW being manipulated with the tools but no  
162 hands were ever seen (thus the bodily movements had to be inferred), while in a third  
163 condition children saw the hands manipulating the tools, but never saw the KW (thus  
164 the mechanical causality needed to be inferred). In a fourth condition, children saw a

165 video of the end-state of the manipulation of the KW, with all the apparatus visible as  
166 it would be after a successful extraction of the reward. These four conditions were  
167 compared to a no-information condition in which children were simply presented with  
168 the KW and no other information, thus allowing a baseline of asocial learning to be  
169 established.

170

---

Figure 1 about here

---

172

173 Like conventional ghost conditions, our ghost control condition removed the  
174 agent from the image, but in a more naturalistic way. In a ghost condition, objects  
175 move in a “ghostly” way that might strike children as rather odd. In our experiment,  
176 children instead saw a tightly focused video view which tracked around the KW to  
177 display the critical aspects of the scene while the agent moving the tools was simply  
178 out of the frame; such images are similar to the kind children often see on television.  
179 For all the video displays the camera moved around the pertinent parts, and children  
180 could see previously completed actions, as well as to be completed actions. We  
181 developed this as a simpler and more natural use of video images than digitally  
182 removing agents, the approach favoured by Huang and Charman (2005). Children as  
183 young as 18-months can imitate behaviour they have seen in televised demonstrations,  
184 even when no narration is presented (Simcock, Garrity, & Barr, 2011), demonstrating  
185 that young children can take information from symbolic media and apply it to real-  
186 world objects (Barr & Hayne, 1999) and by 36 months children’s abilities to imitate  
187 multistep sequences from television demonstrations approaches that of live  
188 demonstrations (McCall, Parke, & Kavanaugh, 1977).

189



190 *Copying hierarchical structure of actions versus style details*

191 As explained above, a child's copying of a series of actions after witnessing a  
192 display is often only partial; thus one can ask what influences which features of the  
193 display are copied? By manipulating the type of information observed we can  
194 establish whether witnessing certain forms of information facilitates the copying of  
195 specific types of behaviour. For example, children who witness displays that focus on  
196 a model's bodily actions, without the same level of functional information as  
197 presented in a ghost control condition, may be more likely to copy action styles (in the  
198 present study, whether a model tapped or twisted a tool) as this is a main focus of  
199 what has been witnessed.

200 The two main aspects of action structure we investigated were the hierarchical,  
201 sequential structure of the actions (Byrne & Russon, 1998; Flynn & Whiten, 2008;  
202 Whiten *et al.*, 2006) and the action "style" (Flynn & Whiten, 2008; Hobson & Lee,  
203 1999). With respect to the first of these, the KW box was designed to allow an  
204 examination of the program-level copying of hierarchical action structure; that is the  
205 "copying the structural organization of a complex process (including the sequence of  
206 stages, subroutine structure, and bimanual coordination), by observation of the  
207 behaviour of another individual, while furnishing the exact details of actions by  
208 individual learning" (Byrne & Russon, 1998, p. 677). In line with previous studies  
209 using the KW, this task allowed an investigation of the imitation of sequential  
210 structuring within actions of appropriate complexity (as in copying a series of acts  
211 A,B,C versus C,B,A); and (ii) imitation of hierarchical structuring, that goes beyond  
212 mere replication of linear sequences, recognizing instead the way in which lower level  
213 elements of behaviour are embedded within a higher-level organization. For  
214 successful extraction of a reward from the KW, sixteen actions could be performed in

215 multiple ways, but in the displays used in the current study the actions were presented  
216 in either of two different hierarchically organised sequences (“hierarchical  
217 sequences”, for short), that we call the “Row” versus “Column” approaches (Flynn &  
218 Whiten 2008; Whiten *et al.*, 2006; see the Methods section for further details).  
219 Presenting these two alternative hierarchical sequences allowed us to establish  
220 whether the different viewing conditions facilitated or inhibited the acquisition of this  
221 type of information. The hierarchical sequence of operation was discernible in all the  
222 conditions except the end-state and no information conditions. We predicted, in line  
223 with ghost control studies in which a whole display produced the most faithful  
224 copying performance, that children who witnessed the whole display would show a  
225 higher level of adoption of the witnessed hierarchical sequential structure than  
226 children who witnessed other viewing conditions. Further, in line with evidence on a  
227 simpler task (Huang & Charman, 2005) we predicted that when the interactions of the  
228 tools and apparatus were presented, more copying of hierarchical actions would occur  
229 than in conditions that presented information about hands and tools only.

230 Our study also considered children’s adoption of “styles” of specific  
231 subsidiary actions, in this case tapping or twisting the tools, based on the viewing  
232 condition witnessed. For this it was appropriate to compare only the whole display  
233 with the hands-and-tools-only condition, as action styles were present only in these  
234 conditions. Bekkering *et al.* (2000) found that, depending on context, different  
235 elements of a display can take precedence. For example, action “style” can become  
236 the focus of a series of actions in contexts in which the end-state is not emphasised: in  
237 conditions in which a child witnessed a demonstration of a mouse hopping and then  
238 being placed inside a house, versus a mouse hopping to the same location with no  
239 house present, children were more likely to reproduce the hopping when the house

240 wasn't there (Carpenter, Call, & Tomasello, 2005). We predicted that children would  
241 be more likely to copy the action style in the condition in which they saw the hands  
242 and tools only, as these action styles would have been a main focus, compared to the  
243 whole display that incorporated multiple goals, including these action styles.

## 244 **Method**

### 245 *Participants*

246 One hundred and forty 3-year-olds and one hundred and forty 5-year-olds  
247 participated in this yoked design study. Children within each age group were matched  
248 across different conditions according to verbal mental age, as measured by the British  
249 Picture Vocabulary Scale (BPVS, Dunn, Dunn, Whetton, & Pintilie, 1997), so that  
250 there was no more than three months difference between them and their yoked  
251 participants in the other conditions. The mean difference between the yoked  
252 participants was one month. Descriptive statistics for the participants in each of the  
253 conditions are shown in Table 1; within each age group comparisons across the  
254 conditions show no significant difference in chronological age, 3-year-olds,  $F(4, 140)$   
255 = .46, *ns.*, 5-year-olds,  $F(4, 140) = .45$ , *ns.*

256

---

257 Table 1 about here

---

258

### 259 *Design*

260 A between-group design was used, in which children were assigned to one of  
261 several experimental conditions or to a no information control. The four main types of  
262 experimental condition differed in the form of information provided in a  
263 demonstration video (see Figure 2): (i) the “whole display” showed the KW, tools and  
264 hands of the demonstrator (presenting full information), (ii) the “box-and-tools-only

265 display” showed only the KW and tools being manipulated and never showed the  
266 demonstrator or the demonstrator’s hands (presenting emulation information), (iii) the  
267 “hands-and-tools-only display” showed the demonstrator’s hands manipulating the  
268 tools, but the KW box was never seen (presenting bodily movement imitation) and  
269 (iv) the “end-state display” condition showed the KW and tools as they would be after  
270 it had been opened, but showed no moves towards this end-state. The final condition  
271 was a no information control condition in which no information was given before a  
272 participant was presented with the KW. For the video displays, the camera tracked  
273 around the hands and/or box so that the pertinent parts of the scene, in line with the  
274 display being observed (e.g., a shape being inserted into the front of the lid or the  
275 tools being tapped), were presented (videos can be viewed in the Supplementary  
276 Material). The displays presented as much of the scene as was possible without  
277 including the part of the display that was to be obscured; thus the displays allowed the  
278 sequence of actions to be seen as children saw the result of a previous action, as well  
279 as the following apparatus to be manipulated (e.g., that there was a space for a tool to  
280 be inserted; or that there was a shape into which a tool was yet to be inserted). Also,  
281 the final scenes of the video reflected the information being presented, with the whole  
282 display and box-and-tools-only display showing the KW open with the keys  
283 assembled (details below) next to the box, while the hands-and-tools-only display  
284 ended after the image showed the model’s hands moving to lift the lid off of the KW.

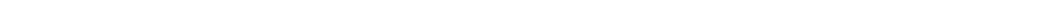
285



286

Figure 2 about here

287



288

289

The whole display, box-and-tools-only and hands-and-tools-only conditions were divided further at two separate levels. For each of these conditions children saw

290 one of two types of display, Row or Column, which differed according to the  
291 hierarchical sequence of the actions undertaken. Both the Row and Column displays  
292 incorporated the same set of twelve operations on the box (the twelve actions are a  
293 result of the sixteen actions minus the four actions on a missing tablet, described  
294 later), but these elements were organised into alternative hierarchical sequences. The  
295 second level of division related to the manner in which the tools were manipulated; in  
296 half of the displays the tools were tapped and in the other half the tools were twisted  
297 into tablets as described below. These differences resulted in fourteen different  
298 conditions for each age group, as illustrated in Table 1. Videos for each of the video  
299 displays can be found in the supplementary material.

### 300 *Materials*

301 The task used in this study, the “Keyway Fruit” (KW, see Figure 1) was  
302 almost completely transparent and explicitly designed to study the imitation of  
303 hierarchically-structured, complex actions sequences (Flynn & Whiten, 2008, Whiten  
304 *et al.*, 2006). A lid was fitted to the box in the manner of a shoe-box lid held in place  
305 by four skewers running through both lid and box. The skewer ends did not protrude  
306 so could not be removed by fingers alone. On top of the lid was a row of four hollows  
307 of different shapes; each hollow contained a different coloured plastic tablet of the  
308 same shape. By stabbing a stick-tool into a hole in each tablet, the tablet could be  
309 lifted up. The “key” thus formed could then be inserted into a correspondingly-shaped  
310 hollow at the front of the lid, thus pushing backwards one of the skewers that  
311 protruded into the hollow. The other end of the skewer could then be grasped and  
312 removed. As an incentive to open the KW, a capsule that contained a reward (a  
313 sticker) was placed inside.

314 *Procedure*

315           Each participant was seen twice. During the first session children were tested  
316 using the BPVS. From this test children’s verbal mental ages were calculated and the  
317 participants were yoked according to these verbal mental age scores across the  
318 conditions. During the second testing session each participant in the experimental  
319 conditions sat next to an experimenter in front of a laptop computer. The child was  
320 told, “You sit here and watch what happens on the computer because I’m going to let  
321 you have a go (pointing to the KW, which was covered with a cardboard box) in a  
322 minute.” The child proceeded to watch one of thirteen video displays, which showed  
323 opening of the KW, or the “end-state display” only.

324           The video displays differed in three ways, (i) the type of information presented  
325 (“whole display”, “box-and-tools-only display”, “hands-and-tools-only display” or  
326 “end-state display”), (ii) the hierarchical order of actions used to open the box (Row  
327 or Column), and (iii) the manner in which the tool was inserted into each tablet  
328 (twisted or tapped). The videos differed from “ghost control” video displays (Huang  
329 & Charman, 2005) that were digitally altered so that the whole scene was presented  
330 but certain elements were removed (for example, a block might be seen to float into a  
331 hole as the model’s hand and arm have been digitally removed). We presented what  
332 we see as a more ecologically-valid display by simply focusing the camera on the  
333 appropriate part, so that only the box and tools could be seen, or only the hands and  
334 tools and not the KW. Such an approach meant that the size of certain elements on the  
335 laptop screen were slightly different for each viewing condition (as can be seen in  
336 Figure 2), but (and as the results will show) seeing a larger version of certain elements  
337 did not necessarily mean that it was more likely to be copied.

338           In the “Column” approach to the task, a key was made with the first tablet and  
339 inserted in the corresponding hollow, then the skewer and the key were removed  
340 (picture b in Figure 1). This “Column” of procedures was then repeated with each  
341 tablet in turn. In the alternative “Row” displays, the actions were completed  
342 consecutively along each row (picture c in Figure 1). Thus, tools were first inserted  
343 into all tablets, then all keys into hollows, after which all skewers were removed,  
344 followed by all keys. Both the Column and Row video displays had two versions, one  
345 which showed the tools being twisted into the holes in the tablets, and one which  
346 showed the tools being tapped into the holes in the tablets, by holding it with one  
347 hand and tapping down on top with the other hand. The Row/Column distinction  
348 allowed hierarchical-sequence imitation to be investigated and the tap/twist distinction  
349 allowed the investigation of the imitation of action style.

350           If a child appeared distracted while watching the video, s/he was told, “Watch  
351 carefully because you are going to have a go in a minute.” No reference was made to  
352 the KW or the goal of extracting the capsule, in line with previous studies which have  
353 used the KW (Flynn & Whiten, 2008; Whiten *et al.* 2006). As with previous studies,  
354 after the display finished, each child was presented with the KW and simply told,  
355 “Now it’s your turn.” If the child did not interact with the KW, or asked for help, s/he  
356 was asked, “What do you think you do? Can you show me?” If the child was still  
357 reluctant to continue the experimenter said, “You’re doing really well. Can you show  
358 me what you think you do next? Walk all the way round the box and have a look to  
359 see if there is anything that you think you do.” After the child’s first attempt, whether  
360 successful or not, each child was told, “Let’s watch the video again and then you can  
361 have another go.” We intended to use a proviso that children would be discounted if  
362 they did not attend to the videos after such prompts; however, it was not necessary to

363 use this rule as all children attended to the videos. The KW was reassembled out of  
364 sight and the procedure began again; thus, all participants in the experimental  
365 conditions attempted to open the KW twice, and both attempts were preceded by a  
366 viewing of the same demonstration.

367         The KW in the video displays differed slightly to the KW in the testing  
368 sessions because in the displays the third tablet (the “T” shaped tablet as shown in  
369 Figure 1) was always absent, but when the child was presented with the KW all the  
370 tablets were present. Introducing the “missing” tablet in this way tested whether a  
371 child was merely copying a chain of actions, in which case the new tablet was  
372 predicted to be left until last or ignored. If a child had acquired the hierarchical  
373 sequence, the new tablet would be assimilated to a hierarchical order and dealt with  
374 during the third position.

375         In the no information control condition each child was shown the KW and the  
376 experimenter said, “Can you see this box? What do you think you do with it? Can you  
377 show me?” Then the child was allowed to interact with the KW. If the child didn’t  
378 interact with the KW the experimenter followed the same series of prompts as in the  
379 experimental conditions.

#### 380 *Coding*

381         All trials were videotaped for later coding. A number of behaviours were of  
382 interest: the number of transitions between consecutive actions (the movement from  
383 manipulation of a pertinent part of the KW to the manipulation of another pertinent  
384 part of the KW) that occurred either along rows or down columns (as described  
385 below), the number of tablets into which tools were tapped or twisted, time taken on  
386 the task (from the child’s first touch of the box to his/her last touch, success or refusal  
387 to participate after prompts) and the position of the manipulation of the tablet that had



388 been missing in the demonstration. Row-wise transitions occurred when a child  
389 progressed from one action to the same kind of action on a similar object, such as  
390 stabbing a tool into one tablet, then stabbing another tool into a further tablet (one row  
391 transition) or removing one skewer, then a second and a third (two row transitions).  
392 Column-wise transitions occurred when a child completed consecutive actions  
393 concerning the same tablet. Thus, stabbing a tool into a hole in a tablet, placing the  
394 key thus made into the front hollow, then removing the key followed by the skewer  
395 consisted of three column-wise transitions. The percentage of row-wise transitions for  
396 each child was calculated by dividing the number of row transitions by the sum of row  
397 and column transitions and then multiplying by 100. Similar calculations were  
398 performed for column transitions and for tapping and twisting (tapping moves divided  
399 by the number of tapping and twisting moves). An index of the imitation of the  
400 hierarchical sequence was calculated by dividing the number of transitions made  
401 which were faithful to the method a child had witnessed by the number of total  
402 transitions made. A similar calculation was used to provide an index of “action style”  
403 imitation. Inter-rater reliability for the sequence of actions for 85 attempts selected at  
404 random (16% of the total attempts) produced a Cohen’s kappa of 0.93 for the Row  
405 moves and 0.95 for the Column moves. Nearly all discrepancies occurred within the  
406 no information control condition.

407

## Results

408 Four questions were investigated: (i) was there evidence of observational  
409 learning and what effect did the different forms of information presentation have on  
410 children’s subsequent levels of success?; (ii) did children copy the hierarchical  
411 sequence they observed?; (iii) did children incorporate a missing tablet, and if so  
412 where?; and (iv) did children copy the action style they witnessed (comparing the

413 whole-display condition and hands-and-tools-only conditions, as no action detail was  
414 observed in the box-and-tools-only condition)? Each question was addressed in  
415 relation to age effects (3- versus 5-year-olds) and the type of information presented.  
416 As there was no effect of gender this is not considered further.

417 Repeated-measures analyses of variance (ANOVA) showed that the main  
418 difference between the children's first and second attempts at the KW was that  
419 children made significantly more transitions in their second attempt ( $M = 5.99$ ) than  
420 their first attempt ( $M = 4.80$ ),  $F(1, 259) = 45.13$ ,  $p < .0001$ , partial  $\eta^2 = .15$ . For  
421 economy the following analysis will concentrate on the first attempt, with the results  
422 of the second attempt only being reported when additional effects were found.

423

424 *Was there evidence of observational learning and how important was the type of*  
425 *information witnessed for a child's level of success?*

426 Figure 3 shows the percentage of children who successfully opened the KW  
427 according to age (3- versus 5-year-olds) and the type of information (whole-display,  
428 hand-and-tools- only, box-and-tools-only, end-state and no information control)  
429 witnessed. Five-year-olds were more successful at opening the KW than 3-year-olds,  
430  $\chi^2(1, N = 280) = 14.55$ ,  $p < .0001$ ; and children who watched the whole-display were  
431 significantly more successful than children in all the other conditions,  $\chi^2$  ranged from  
432 10.43 to 35.34,  $N$  ranged from 100 to 160, all  $ps < .001$ . Children in the box-and-  
433 tools-only condition were significantly more successful than children in the hands-  
434 and-tools-only condition,  $\chi^2(1, N = 160) = 6.94$ ,  $p < .05$ . There were no other  
435 significant differences.

436 It is important to compare the behaviour of children in the no information  
437 control and end-state conditions with the children in the experimental conditions at

438 this point, as it is not possible to include these children in some of the future analyses  
439 because they did not witness a demonstration. There was a significant effect for the  
440 type of information a child witnessed (whole-display, hand-and-tools- only, box-and-  
441 tools-only, end-state and no information control) in terms of the number of moves  
442 made, ANOVA,  $F(4, 280) = 10.50, p < .0001$ , partial  $\eta^2 = .13$ . Children in the whole-  
443 display condition made significantly more moves ( $M = 6.33$ ) than children in the  
444 hands-and-tools-only ( $M = 3.41$ ), end-state ( $M = 2.85$ ) and no information conditions  
445 ( $M = 3.90$ ). Children in the box-and-tools-only condition ( $M = 5.16$ ) made  
446 significantly more moves in their first attempt than children in the hands-and-tools-  
447 only and end-state conditions (all Bonferonni post hoc tests,  $p < 0.05$ ). Importantly,  
448 children in the no information condition spent significantly longer manipulating the  
449 KW ( $M = 6$  minutes) than children in the end-state ( $M = 3$  minutes) and experimental  
450 conditions ( $M = 3\frac{1}{2}$  minutes), ANOVA,  $F(2, 267) = 27.16 p < .001$ , partial  $\eta^2 = .17$ ,  
451 Bonferonni post hoc  $p < 0.05$ ; suggesting that the difference in the number of moves  
452 was not due to a lack of interaction with the KW.

453

---

Figure 3 about here

---

454

455

456

457 *Did children copy the hierarchical sequence they observed?*

458 Children who witnessed the Row demonstrations made significantly more row  
459 moves ( $M = 4.08$ ) than children who witnessed the Column demonstrations ( $M =$   
460  $1.87$ ), ANOVA,  $F(1, 239) = 30.97, p < .001$ , partial  $\eta^2 = .12$ . Similarly, children who  
461 witnessed the Column demonstration made significantly more column moves ( $M =$   
462  $2.81$ ) than children who witnessed the Row demonstration ( $M = 1.18$ ), ANOVA,  $F(1,$

463 239) = 20.16,  $p < .001$ , partial  $\eta^2 = .08$ . Note that this second effect does not follow  
464 automatically from the first because children can perform moves other than Row and  
465 Column ones, for example, moving diagonally across the KW to manipulate a  
466 different piece.

467 Two levels of analysis for the imitation of the hierarchical sequence were  
468 carried out: (i) the absolute number of moves that were in line with the hierarchical  
469 order witnessed, and (ii) the percentage of moves that were in line with the  
470 hierarchical order witnessed. The initial analysis of the absolute number of moves  
471 made in line with the hierarchical sequence witnessed showed an effect for age,  
472 MANOVA,  $F(1, 239) = 19.61$ ,  $p < .001$ , partial  $\eta^2 = .08$ , and the type of information  
473 witnessed (whole-display, hand-and-tools- only and box-and-tools-only),  $F(2, 239) =$   
474  $16.21$ ,  $p < .001$ , partial  $\eta^2 = .12$ . Five-year-olds made significantly more moves in line  
475 with the hierarchical order witnessed ( $M = 4.37$ ) than 3-year-olds ( $M = 2.51$ ).  
476 Children who witnessed the whole-display made more moves in keeping with the  
477 hierarchical sequence they witnessed ( $M = 4.99$ ) than children who witnessed the  
478 box-and-tools-only display ( $M = 3.28$ ) or children who witnessed the hands-and-  
479 tools-only display ( $M = 2.06$ ; all Bonferroni post hoc tests  $p < .05$ ). The second  
480 attempt produced the same effects, and in addition, children who witnessed the box-  
481 and-tools-only display made significantly more moves in keeping with the  
482 demonstration they had witnessed ( $M = 4.61$ ) than children who witnessed the hands-  
483 and-tools-only display ( $M = 2.59$ ), Bonferroni post hoc tests  $p < .05$ .

484 A measure of the percentage of moves in line with the hierarchical sequence  
485 witnessed was calculated (number of moves made which corresponded to the  
486 demonstration witnessed divided by the total number of row and column moves made,  
487 and multiplied by 100). There was no age effect for percentage of moves in line with

488 the hierarchical sequence witnessed, MANOVA,  $F(1, 205) = .83$ , *ns*. An effect for  
489 type of information (whole-display, hand-and-tools- only and box-and-tools-only)  
490 witnessed approached significance at the initial attempt,  $M$  (whole-display) = 74%  
491 and  $M$  (box-and-tools-only) = 57%,  $F(2, 205) = 3.03$ ,  $p = .051$ . At the second attempt,  
492 the lack of effect for the type of information remained, although it approached  
493 significance, MANOVA,  $F(2, 219) = 2.74$ ,  $p = .07$ , but an age effect was found with 5-  
494 year-olds producing a higher percentage of hierarchical imitative moves ( $M = 73\%$ )  
495 than 3-year-olds ( $M = 59\%$ ),  $F(1, 219) = 7.49$ ,  $p < 0.01$ , partial  $\eta^2 = .03$ ; although this  
496 effect was weak.

497

498 *Did children incorporate the missing tablet into a hierarchically-organised*  
499 *approach?*

500 Figure 4 illustrates the position in which children dealt with the missing tablet,  
501 allowing an investigation of whether children incorporated this new aspect of the task  
502 into a hierarchical sequence. 40% of the experimental groups ( whole-display, box-  
503 and-tools-only and hands-and-tools-only) incorporated it at the “correct” third  
504 position, showing an effect significantly above chance (Binomial test with a test  
505 proportion of .25 as the likelihood of incorporating the piece at any of the four  
506 positions,  $n = 280$ ,  $p < .001$ ). Also, both 3- and 5-year-olds were more likely than  
507 chance to incorporate the third tablet at the “correct” third position, 3-year-olds,  
508 Binomial test,  $n = 140$ ,  $p < .05$ ; 5-year-olds, Binomial test,  $n = 140$ ,  $p < .05$ .

509

510

---

Figure 4 about here

---

511

512 *Did children copy the action styles they witnessed?*

513 In order to establish whether there was imitation of style details, children's  
514 tool tapping or twisting manipulations were coded. At the first attempt only 16% of  
515 the children who witnessed the tools being tapped into the holes copied this tap, and  
516 40% of the children who witnessed the tools being twisted into the holes copied the  
517 twist. Although less than half of children produced an action styles, when they did so  
518 it was overwhelming faithful to the action style witnessed, rather than adhering to a  
519 predisposed action style. Children who had witnessed the tapping, performed tapping,  
520  $\chi^2(1, n = 160) = 11.27, p < .01$ , and children who had witnessed the twisting,  
521 performed twisting,  $\chi^2(1, n = 160) = 23.33, p < .001$ . There was no difference between  
522 the age groups (3- versus 5-year-olds) in the imitation of detailed actions,  $\chi^2(1, n =$   
523  $160) = .77, ns.$ : 29% and 38% of 3-year-olds imitated detailed actions during the first  
524 and second attempts respectively, while 25% and 43% of 5-year-olds imitated detailed  
525 actions during the first and second attempts respectively. Similarly there was no effect  
526 for the type of information (whole-display and hand-and-tools- only) witnessed on the  
527 level of detail imitated,  $\chi^2(1, n = 160) = .74, ns.$ : 31% and 45% of children who  
528 witnessed the "whole-display" imitated detailed actions during the first and second  
529 attempts respectively, while 25% and 35% of children who witnessed the "hands-and-  
530 tools-only" display imitated detailed actions during the first and second attempts  
531 respectively. There was no association between children's ability to open the KW and  
532 their adoption of the detailed actions for the different age groups, 3-year-olds,  $\chi^2(1, n$   
533  $= 80) = .41, ns.$ ; 5-year-olds,  $\chi^2(1, n = 80) = 1.17, ns.$ , or in the type of information  
534 witnessed, whole-display,  $\chi^2(1, n = 80) = .27, ns.$ ; hands-and-tools-only,  $\chi^2(1, n = 80)$   
535  $= .12, ns.$

536

537

## Discussion

538           Our aim was to establish the role of different aspects of what children witness  
539 during witnessing completion of a complex task on their subsequent imitation of  
540 hierarchical-sequence information and action style, as well as their success at  
541 completing the task. Importantly, our new approach to video manipulation  
542 successfully allowed us to address these questions.

543

### 544 *Dissecting imitation versus emulation*

545           Considering the level of success alone, children who witnessed a display  
546 which contained all the information were significantly more successful at opening the  
547 KW than children who witnessed other displays; and children who witnessed the box-  
548 and-tools-only information (which provided emulative information about how the  
549 objects could be moved, but lacked information about the bodily movements  
550 necessary to make such movements) were significantly more successful at opening the  
551 KW than children who witnessed hands-and-tools-only information (which provided  
552 imitative information about how the hands could be moved in relation to the tools, but  
553 lacked the emulative information about how the tools and box interacted). Such  
554 findings are in line with previous research with younger children on simpler tasks, and  
555 extends this work to more a more complex task (Hopper, *et al.* 2010; Tennie, *et al.*  
556 2006). Thus in terms of general success none of our manipulated forms of information  
557 provided the same level of support to the learner as that witnessed in the whole  
558 display, but when comparing manipulated displays, emulative information regarding  
559 object manipulations was more useful for success on the KW than information  
560 regarding the imitation of hand movements.

561 As well as reduced success in extraction of the reward from the KW for  
562 children who witnessed the hands-and-tools-only display (presenting information  
563 about a model's hand movements in relation to tool manipulations, but lacking  
564 information about how the objects interacted), there was also a reduction in the  
565 copying of hierarchical-sequence information. Such a finding is consistent with the  
566 fact that although these children witnessed the spatial sequence of the actions, it was  
567 less clear how this sequence of actions related to the KW and successful extraction of  
568 the reward. In contrast, the lack of fidelity to the hierarchical sequence and the  
569 reduced level of success for children who witnessed displays that contained box-and-  
570 tools-only (emulative) information in comparison to the whole displays is surprising,  
571 as the effect of the sequence of tool manipulations in relation to the KW in the box-  
572 and-tools-only displays was visible, along with information about the previous and  
573 subsequent actions. The only element that was lacking from this display was the  
574 model's hands and the hand movements. Thus, it appears that displays that present  
575 information regarding how objects interact with one another in a complex task (but  
576 don't display a model's hands), are not equivalent to a whole display for children as  
577 young as those we studied.

578 A number of possibilities might explain this significant finding. First, it could  
579 be that the difference in size of the elements in the video presentations caused the  
580 reduction of hierarchical sequence information reproduction. As Figure 2 shows, the  
581 display containing the box-and-tools-only information provided a slightly larger  
582 version of the KW, while the whole display contained the same information and also  
583 the model's hands and so, as a result the KW is slightly smaller on the screen. For  
584 both displays the camera tracked to focus on the workings of the pertinent part of  
585 interest, as well as aspects within the immediate frame (previously manipulated item,



586 and subsequent action item). It seems unlikely that having a larger view of the point at  
587 which specific information is given, in this case hierarchical sequence information,  
588 provides less information to an observer, and we do not believe that the different size  
589 of the KW in each of the displays is the cause of the disparity in the adoption of  
590 hierarchical sequence information or success.

591         The critical distinction between the two displays was that one contained the  
592 model's hands, the tools and the KW, while the other only contained the tools and the  
593 KW. Perhaps seeing the model's hands provided children with some reference about  
594 where they should put their hands in relation to the tools. Alternatively, having a  
595 "social" element, another person (even just their hands), within the display may  
596 facilitate greater focus on the display, as recorded for 24 month-olds (Slaughter,  
597 Nielsen, & Enchelmaier, 2008). Future work could address the importance of the  
598 presence of the model's hands as a social aspect of the scene, perhaps by having  
599 hands in the scene that are not active. Hopper *et al.* (2008) used an enhanced ghost  
600 control, as described in our introduction, albeit with a much simpler task in which a  
601 door was moved left or right using fishing wire, with a non-active child sitting  
602 passively waiting for the reward to exit the box present within the scene. She found  
603 that children copied the movement of the door in their first response for both a full  
604 demonstration and an enhanced control condition; but across all trials a higher level of  
605 matching occurred after the full demonstration compared to the enhanced ghost  
606 condition, with the level of matching being similar in both the standard ghost control  
607 that did not contain an additional social element and the enhanced ghost control. In  
608 contrast to the KW task used in the current study, Hopper *et al.* (2008) used a very  
609 simple task illustrated by the fact that six out of eight children who were presented  
610 with the task in a no information condition were able to successfully extract the

611 reward, a level of success that did not significantly differ to the success of the full  
612 demonstration group. Thus it remains unclear whether attempting an enhanced ghost  
613 control condition by adding a set of passive hands in the current study's box-and-  
614 tools-only condition with our more complex task, thus signalling a social aspect to the  
615 demonstration but providing no information about how these hands move the tools,  
616 would result in more initial copying but less fidelity over trials (as in Hopper et al.  
617 2008) or whether the greater complexity of the KW would produce different results.

618

#### 619 *Copying hierarchical structure of actions versus style details*

620 Overall, children showed fidelity to whichever of the two hierarchical action  
621 sequences they witnessed. A developmental change was seen in the replication of the  
622 hierarchical sequences, as 5-year-olds showed a stronger tendency than did 3-year-  
623 olds to adopt the witnessed hierarchical sequence. Such a difference may be  
624 explicable by 5-year-olds' more advanced memory and/or cognitive skills.

625 Access to information on actions, tool operations and the object affected also  
626 influenced the adoption of hierarchical sequence information; children who witnessed  
627 the whole display made significantly more moves in line with the hierarchical  
628 sequence information that they witnessed than children who witnessed more limited  
629 displays. Interestingly, by the children's second attempt, those who had witnessed the  
630 tools-and-box-only display (which provided emulative information about how the  
631 objects could be moved, but lacked information about the bodily movements  
632 necessary to make such movements) reproduced significantly more actions in line  
633 with the hierarchical sequence information that they had witnessed than children who  
634 had seen the hands-and-tools-only display (which provided imitative information  
635 about how the hands could be moved in relation to the tools, but lacked the emulative

636 information about how the tools and box interacted). It appears that with some  
637 personal experience with the task during their initial attempt and then a subsequent  
638 demonstration, children were able to discern and reproduce more of the hierarchical  
639 sequence information, but this was true only when the witnessed display contained  
640 functional information about the KW (the box-and-tools-only display), and not when  
641 it contained information about bodily movements and tool manipulations (the hands-  
642 and-tools-only display).

643         Comparing the whole display and the hands-and-tools-only display allowed  
644 the reproduction of the different action styles, tapping versus twisting the tools, to be  
645 investigated. It was predicted that as the hands-and-tools-only displays contained little  
646 information beyond how the tools were manipulated, there should be significantly  
647 more replication of these action styles in this condition than in the whole display,  
648 which contained a much richer and more varied series of information about a number  
649 of different goals including a clearer demonstration of the hierarchical sequence  
650 information. Carpenter *et al.* (2005) and Bekkering *et al.* (2000) showed that when  
651 there are competing goals, action style will not be replicated at the expense of an end-  
652 state. In the current study, of those children who did reproduce either tapping or  
653 twisting, there was fidelity to the method that they witnessed. However, this  
654 replication of action style was not dictated by either age or, more surprisingly, by the  
655 type of information that a child had witnessed. That is, children who witnessed the  
656 hands-and-tools-only information were not more likely to imitate the action styles  
657 than children who witnessed the whole display. It was also not the case that children  
658 who copied the action styles did so at the expense of learning more about how to open  
659 the KW; there was no relation between copying the action style and the successful  
660 retrieval of the reward from the KW.

661           The lack of a relation between the number of goals witnessed in a display (in  
662 this case the whole display contained more goals than the hands-and-tools-only  
663 display) and the replication of the action style witnessed is difficult to explain. It may  
664 be that the complexity of the task presented to the children in the current study  
665 affected their replication of the action style, as previous studies in relation to the  
666 organisation of the replication of goals have used simpler tasks (moving a toy mouse  
667 along a path and placing it in a house, Carpenter *et al.*. 2005; touching a dot on a table  
668 with one's hand, Bekkering *et al.*, 2000). However, such an explanation does not  
669 seem logical, insofar as the KW was a more complex task, and thus had more sub-  
670 goals for the children who watched the whole display, while those who watched the  
671 hands-and-tools-only display will have seen fewer sub-goals, and so could have been  
672 expected to replicate the action style more; but this did not happen. Alternatively, it  
673 may have been that children understood the affordance of the KW and the tools, such  
674 that although children in the hands-and-tools-only condition saw the tools being either  
675 tapped or twisted, they appreciated, once presented with the KW, that the tools could  
676 be inserted into the holes in the tablets without either tapping or twisting, and that the  
677 ultimate goal was to insert the tools into the holes in the tablets. If this is the case, this  
678 study suggests that the imitation of action styles is not only flexible in relation to the  
679 goals witnessed within a series of actions, but also in terms of the children's  
680 understanding of the affordances of elements within the task.

681

682

### **Conclusions**

683

684

685

Our aim was to establish the relative importance of different aspects of a display during young children's observational learning of a complex task. We focused on children's acquisition of information with regard to three different aspects of the

686 display: the goal of removing the reward from the KW, the hierarchical structure of  
687 the sequence of actions the model applied, and their action style. Children who  
688 witnessed a display containing all the information in a scene were significantly more  
689 likely to extract the reward from the KW, and to produce hierarchical sequences of  
690 actions, than children who witnessed either box-and-tools-only (which provided  
691 emulative information about how the objects could be moved, but lacked information  
692 about the bodily movements necessary to make such movements) or hands-and-tools-  
693 only information (which provided imitative information about how the hands could be  
694 moved in relation to the tools, but lacked the emulative information about how the  
695 tools and box interacted), or who witnessed only information about the end-state or no  
696 information. A similar pattern of results was produced for the acquisition of  
697 hierarchical sequence information, with children who witnessed the whole display  
698 acquiring hierarchical sequence information more quickly than those who witnessed  
699 either box-and-tools-only information or hands-and-tools-only information. Finally,  
700 the replication of the action styles witnessed was not predicted by the information a  
701 child witnessed, or the child's age. Thus, our overall conclusion is that young  
702 children's observational learning draws simultaneously on several different sources of  
703 information that they witness within a scene; and even when much functional  
704 information is present, as in demonstrations of the objects interacting in our KW task,  
705 observers cannot always successfully extract a reward or replicate the hierarchical  
706 sequence of a series of actions. Although there was a trend for five-year-olds to cope  
707 with partial information better than the three-year-olds (as can be seen in Figure 3)  
708 this failed to achieve statistical significance. Whether multiple sources of information  
709 gain more redundancy for older children remains to be systematically studied: for  
710 another quite different task (making a paper aeroplane that flies well, Caldwell &

711 Millen, 2009) this has been shown to be the case for young adults. However it appears  
712 that in observational learning of children as young as those we studied faced with a  
713 complex task, as in many other parts of psychology, “the whole is greater than the  
714 sum of its parts”.

715

716 **References**

- 717 Akins, C. K., Klein, E. D. & Zentall, T. R. (2002). Imitative learning in Japanese quail  
718 (*Coturnix japonica*) using the bidirectional control procedure. *Animal*  
719 *Learning and Behavior* 30, 275–281.
- 720 Barr, R. & Hayne, H. (1999). Developmental changes in imitation from television  
721 during infancy. *Child Development*, 70, 1067-1081.
- 722 Bekkering, H., Wohlschläger, A., & Gattis, M. (2000). Imitation of gestures in  
723 children is goal-directed. *Quarterly Journal of Experimental Psychology*, 53,  
724 153–164.
- 725 Byrne, R. W. (1998). Comment on Boesch, C. and Tomasello, M. Chimpanzee and  
726 human culture. *Current Anthropology* 39, 604–605.
- 727 Byrne, R. W., & Russon, A.E. (1998). Learning by imitation: a hierarchical approach.  
728 *Behavioral and Brain Sciences*, 21, 667-721.
- 729 Caldwell, C. & Millen, A. (2009). Social learning mechanisms and cumulative  
730 cultural evolution: is imitation necessary? *Psychological Science*, 20, 1478-  
731 1483.
- 732 Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen- to 18-month-old  
733 infants differentially imitate intentional and accidental actions. *Infant Behavior*  
734 *and Development*, 21, 315-330.
- 735 Carpenter, M., Call, J., & Tomasello, M. (2005). Twelve- and 18-month-olds copy  
736 actions in terms of goals. *Developmental Science*, 8, F13-F20.
- 737 Custance, D. M., Whiten, A., & Fredman, T. (1999). Social learning of an artificial  
738 fruit task in capuchin monkeys (*Cebus apella*). *Journal of Comparative*  
739 *Psychology*, 113, 13-23.

740 Dawson, B. V. & Foss, B. M. (1965). Observational learning in budgerigars. *Animal*  
741 *Behaviour*, 13, 470–474.

742 Dunn, L., Dunn, L., Whetton, C., & Pintilie, D. (1997). *British Picture Vocabulary*  
743 *Scale*. Windsor, U.K.: NFER-Nelson.

744 Fawcett, T. W., Skinner, A. M. J. & Goldsmith A. R. (2002). A test of imitative  
745 learning in starling using two-action method with an enhanced ghost control.  
746 *Animal Behaviour*, 64, 547–556.

747 Flynn, E. (2008). Investigating children as cultural magnets: Do young children  
748 transmit redundant information along diffusion chains? *Philosophical*  
749 *Transactions of the Royal Society B*, 363, 3541–3551.

750 Flynn, E., & Whiten, A., (2008). Imitation of hierarchical structure versus component  
751 details of complex actions by 3- and 5-year-olds. *Journal of Experimental Child*  
752 *Psychology*, 101, 228-240.

753 Hobson, R. P., & Lee, A. (1999). Imitation and identification in autism. *Journal of*  
754 *Child Psychology and Psychiatry*, 40, 649-659.

755 Hopper, L. (2010). 'Ghost' experiments and the dissection of social learning in  
756 humans and animals. *Biological Reviews*. 85, 685-701.

757 Hopper, L., Lambeth, S., Schapiro, S., & Whiten, A. (2008). Observational learning in  
758 chimpanzees and children studied through 'ghost' conditions. *Proceedings of the*  
759 *Royal Society: B*, 275, 835-840.

760 Hopper, L., Flynn, E., Wood, L. & Whiten, A. (2010). Observational learning of tool  
761 use in children: Investigating cultural spread through diffusion chains and  
762 learning mechanisms through ghost displays. *Journal of Experimental Child*  
763 *Psychology*, 106, 82-97.



764 Huang, C.-T., & Charman, T. (2005). Gradations of emulation learning in infants'  
765 imitation of actions on objects. *Journal of Experimental Child Psychology*, *92*,  
766 276-302.

767 Klein, E. D. & Zentall, T. R. (2003). Imitation and affordance learning by pigeons  
768 (*Columba livia*). *Journal of Comparative Psychology*, *117*, 414–419.

769 McCall, R., Parke, R., & Kavanaugh, R. (1977). Imitation of live and televised  
770 models by children one to three years of age. *Monographs of the Society for*  
771 *Research on Child Development*, *42* (5, Serial No. 173).

772 McGuigan, N., Whiten, A., Flynn, E. & Horner, V. (2007). Imitation of causally-  
773 opaque versus causally-transparent tool use by 3- and 5-year-old children,  
774 *Cognitive Development*, *22*, 353-364.

775 Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of  
776 intended acts by 18-month-old children. *Developmental Psychology*, *31*, 838–  
777 850.

778 Nielsen, M., & Tomaselli, K. (2010). Over-imitation in Kalahari Bushman children  
779 and the origins of human cultural cognition. *Psychological Science*, *21*, 729-  
780 736.

781 Simcock, G., Garrity, K. & Barr, R. (2011). The effect of narrative cues on infants'  
782 imitation from television and picture books. *Child Development*, *82*, 1607-1619.

783 Slaughter, V., Nielsen, M. & Enchelmaier, P. (2008). Interacting socially with human  
784 hands at 24 months of age. *Infancy*, *13*, 185-195.

785 Tennie, C., Call, J. & Tomasello, M. (2006). Push or pull: imitation versus emulation  
786 in human children and great apes. *Ethology*, *112*, 1159-1169.

787 Thompson, D.E., & Russell, J. (2004). The ghost condition: imitations versus  
788 emulation in young children's observational learning. *Developmental*  
789 *Psychology*, *40*, 882-889.

790 Tomasello, M., Kruger, A.C., & Ratner, H.H. (1993). Cultural learning. *Behavioral &*  
791 *Brain Sciences*, *16*, 495-552.

792 Want, S.C. & Harris, P.L. (2002). How do children ape? Applying concepts from the  
793 study of non-human primates to the developmental study of 'imitation' in  
794 children. *Developmental Science*, *5*, 1-13.

795 Whiten, A. (1998). Imitation of the sequential structure of actions by chimpanzees  
796 (*Pan troglodytes*). *Journal of Comparative Psychology*, *112*, 270-281.

797 Whiten, A., Flynn, E., Brown, K., & Lee, K. (2006). Imitation of hierarchical  
798 structure in actions by young children. *Developmental Science*, *9*, 574-582.

799 Whiten, A., & Ham, R. (1992). On the nature and evolution of imitation in the animal  
800 kingdom: Reappraisal of a century of research. In P. J. B. Slater, J. S.  
801 Rosenblatt, C. Beer & M. Milinski (Eds.), *Advances in the Study of Behaviour*,  
802 *21*, (pp. 239-283). New York: Academic Press.

803 Whiten, A, Horner, V., Litchfield, C., & Marshall-Pescini, S. (2004). How do Apes  
804 Ape? *Learning and Behaviour*, *32*, 36-52.

805 Wood, L., Kendal, R., & Flynn, E. (2012). Context dependent model-based biases in  
806 cultural transmission: Children's imitation is affected by model age over  
807 model knowledgeable state. *Evolution and Human Behaviour*, *33*, 387-394.

Table 1. Descriptive statistics for the participants in the different conditions.

	Whole-display				Box-and-tools-only				Hands-and-tools-only				End-State	No Info
	Row		Column		Row		Column		Row		Column			
	Tap	Twist	Tap	Twist	Tap	Twist	Tap	Twist	Tap	Twist	Tap	Twist		
	n=10	n=10	n=10	n=10	n=10	n=10	n=10	n=10	n=10	n=10	n=10	n=10	n=10	n=10
3-year-olds														
Mean Age	44	42	42	43	41	43	43	42	42	42	41	43	42	41
SD	3.39	2.55	3.13	2.71	2.91	3.03	2.75	3.38	3.27	3.10	3.61	3.02	3.37	2.16
Female:Male	6:4	6:4	4:6	3:7	3:7	7:3	4:6	7:3	6:4	4:6	5:5	3:7	4:6	5:5
Mean VMA	38	38	38	38	38	38	38	39	38	39	38	39	38	38
SD	4.64	4.99	6.04	5.53	4.55	5.49	4.66	5.14	4.84	5.04	5.76	5.36	5.21	4.63
5-year-olds														
Mean Age	66	65	65	64	64	64	64	66	65	64	66	65	64	65
SD	3.34	3.96	3.82	3.31	3.28	3.17	2.88	3.33	3.84	3.59	4.22	3.18	3.31	3.08
Female:Male	5:5	3:7	3:7	3:7	5:5	8:2	8:2	4:6	5:5	2:8	6:4	7:3	7:3	5:5
Mean VMA	63	64	64	64	64	64	63	64	64	64	64	64	64	64
SD	12.25	13.40	12.20	12.64	13.12	12.85	12.37	12.78	12.41	12.54	13.01	12.04	12.09	12.94

Note. Chronological age and verbal mental age are presented in months, SD standard deviation, VMA Verbal Mental Age

## Figure Captions:

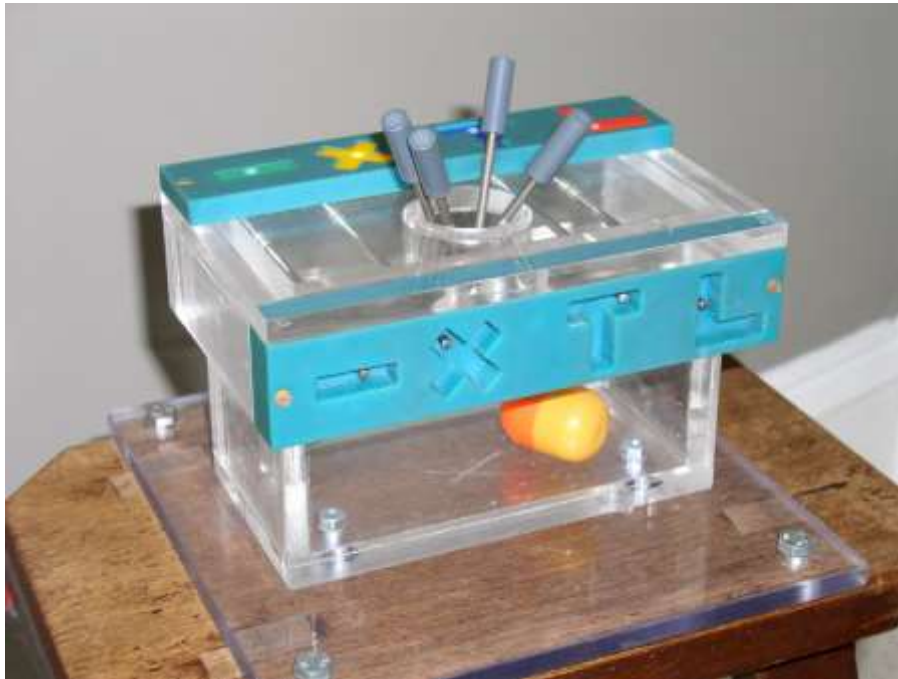
Figure 1. Keyway fruit (a) as presented at the beginning of testing, with the lid held in place by four skewers running through the lid and box (one of the skewers can be seen most clearly in c). In order to open the KW, tools need to be inserted into holes in a series of plastic shapes on the upper face of the KW (as can be seen in b). These “keys” can then be inserted into a series of similarly-shaped hollows at the front-face of the lid (seen most clearly in c). As a result of inserting the keys into the front hollows the skewers, which are holding the lid in place, move through the back of the lid allowing them to be pulled out with one’s hands. The “keys” can then be removed. Once this sequence of actions has been completed with all the KW shapes the lid can be taken off, allowing the reward to be retrieved. This series of actions could be undertaken in a Column-wise sequence, in which all the actions were performed on one Perspex shape, as partly shown in (b), in which the tool is inserted into a hole in a shape, that shape is placed in the hollow, the corresponding skewer is removed, and the key removed from the hollow or in a Row-wise sequence, in which all of the same type of action were performed together, e.g., all the tools are inserted into the holes as shown in (c) and then all shapes inserted into hollows.

Figure 2. Stills from the video displays, (a) whole display, (b) box-and-tools-only display, (c) hands-and-tools-only display, and (d) the end-state display.

Figure 3. Rate of successful extraction of the reward from the KW across age and experimental condition, whole display (WD), box-and-tools-only (B&T), hands-and-tools-only (H&T), end-state (ES), no information (No Info). Note, Children in the no information condition did not have a second attempt.

Figure 4. Number of children manipulating the missing tablet at different positions in their sequence of actions.

Figure 1



(a)



(b)



(c)

Figure 2

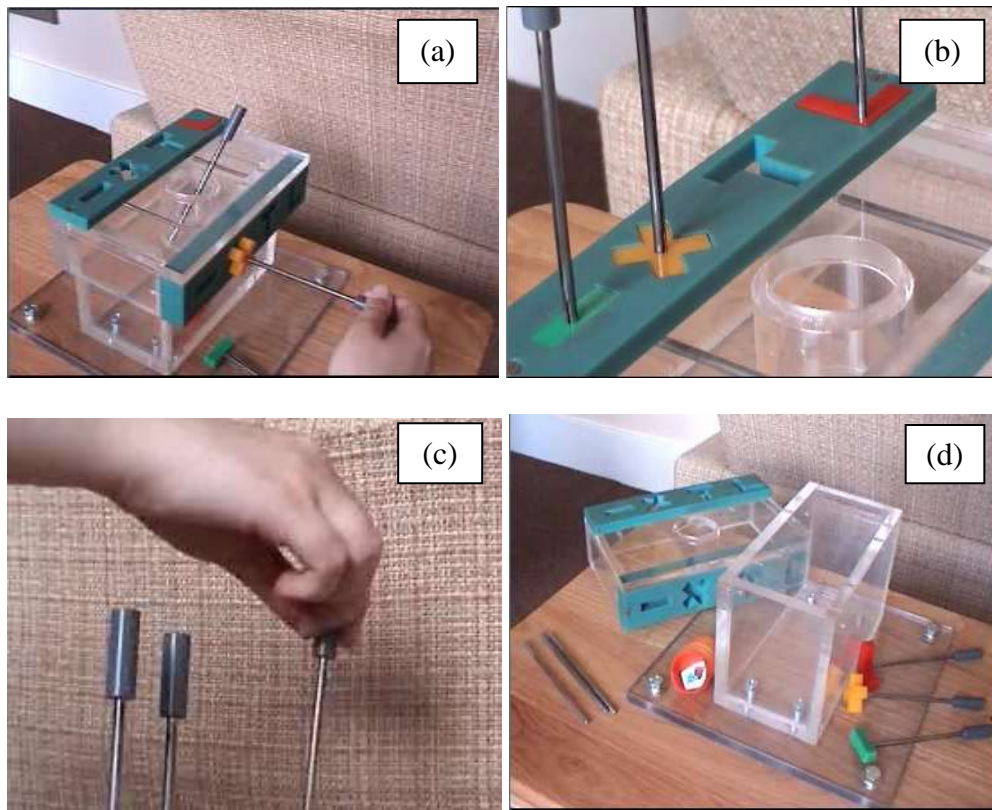


Figure 3

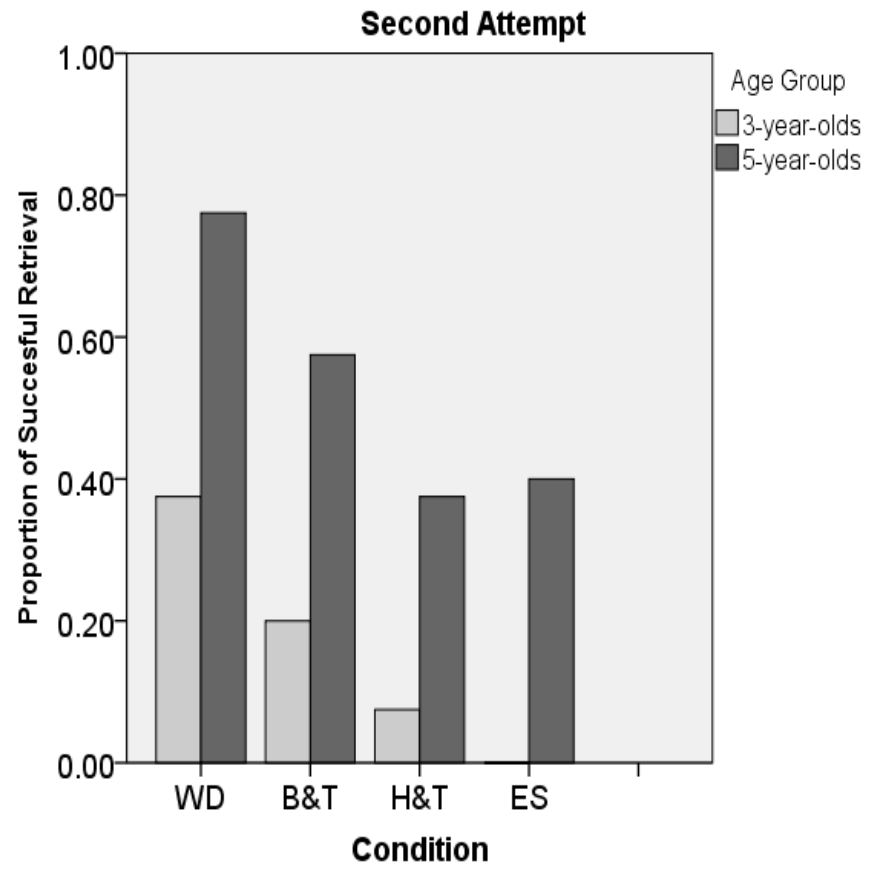
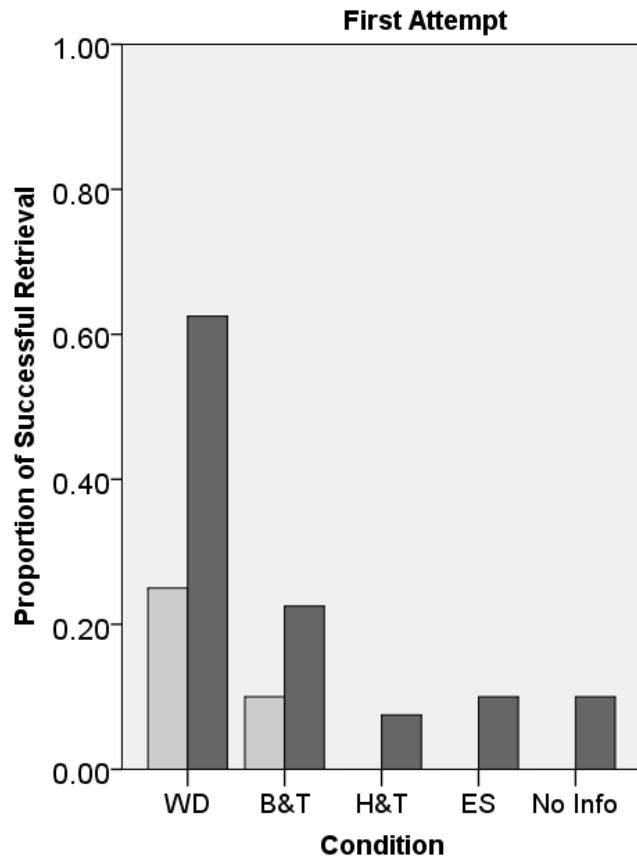


Figure 4

