doi: 10.1111/j.1467-9507.2007.00453.x

Cultural Transmission of Tool Use in Young Children: A Diffusion Chain Study

Emma Flynn, *Durham University*, and Andrew Whiten, *University of St Andrews*

Abstract

Developmental and gender effects in the transmission of information about a tool-use task were investigated within a 'diffusion chain' design. One hundred and twenty-seven children (65 three-year-olds and 62 five-year-olds) participated. Eighty children took part in diffusion chains in which consecutive children in chains of five witnessed two attempts on a tool-use task by the previous child in the chain. Comparisons were made between two experimental conditions in which alternative techniques were seeded and a third no-model control condition. Children in the diffusion chains conformed to the technique they witnessed, in one experimental condition faithfully transmitting a technique absent in the no-model condition. Five-year-olds displayed more robust transmission than three-year-olds, and boys were both more competent and displayed stronger transmission than girls.

Keywords: culture; observational learning; tool use; transmission

Introduction

Culture is one of the key features, along with genetic predisposition and individual learning, which explains our behavior. Culture explains not only much of what we do but, specifically, how we do it. However, unlike genetics or individual learning, the transmission of traditions, which lies at the heart of culture, has received little explicit interest in child psychology, with the notable exception of language acquisition. In particular, no study has addressed the diffusion of traditions across multiple transmission events—so essential for culture—with experimental rigour. Such an oversight is surprising for at least two reasons. Firstly, everyday observation suggests that children are veritable 'culture magnets', picking up the traditions that surround them from a very early age. Secondly, two relatively simple methodologies exist that allow the transmission of cultural variations to be examined in an experimental setting; 'diffusion chain' and 'open diffusion' paradigms (Bartlett, 1932).

In one variant, the diffusion chain, a model (individual A) is trained by an experimenter to perform a behaviour, such as how to open a puzzle box in a particular way. A second individual (B) is then introduced while individual A performs the behaviour.

Correspondence should be addressed to Emma Flynn, Department of Psychology. Durham University, Durham DH1 3LE. Email: *e.g.flynn@durham.ac.uk*

Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

Individual A is then removed and individual B performs while a third individual (C) is present, and so on. Chains continue until the number of participants is exhausted or until the information fails to be transmitted (Horner, Whiten, Flynn, & de Waal, 2006). The other variant, open diffusion, introduces a trained model demonstrating a specific behaviour within a group setting (Whiten, Horner, & de Waal, 2005). Although highly simplified compared with real human culture, open diffusion and diffusion chains offer experimental microcosms that allow the exploration of the transmission of information through observation across 'cultural generations', that is, from one individual to another, affording a high degree of experimental control together with a reasonably ecologically valid environment in which participants learn from each other rather than from an experimenter. 'Cultural generations' thus refer to consecutive social transmissions from individual to individual, which in real life may coincide with genetic generations (parent to child) or not (e.g., peer to peer).

The lack of experimental analyses of culture transmission in children is even more striking when the investigation of culture transmission in human adults and other animals is reviewed. Diffusion studies have examined the transmission of information in human adults (Bangerter, 2000; Mesoudi & Whiten, 2004), communities of chimpanzees (Menzel, Devenport, & Rogers, 1972; Whiten et al., 2005), rats (Laland & Plotkin, 1990), guppies (Reader & Laland, 2000), blackbirds (Curio, Ulrich, & Vieth, 1978), and pigeons (Lefebvre, 1986). Yet, to our knowledge, we are the first to apply this approach to children (Horner et al., 2006).

One key means of cultural transmission is assumed to be observational learning, that is, learning by watching the behaviour of others. Indeed, authorities in developmental psychology (Tomasello, 1999), evolutionary anthropology (Boyd & Richerson, 1985; Richerson & Boyd, 2005), and evolutionary psychology (Plotkin, 2003) have concurred in the view that the evolutionary flowering of human culture was made possible by the emergence of the particularly sophisticated imitative powers that characterize human children. Children's observational learning has been much researched, from the early, influential work of Bandura, Ross, and Ross (1961) to a growing contemporary corpus of studies in which the primary focus is the psychological processes through which a child learns by observation (Gergely, Bekkering, & Kiraly, 2001; Huang & Charman, 2005; Huang, Heyes, & Charman, 2002; Hurley & Chater, 2005; Meltzoff & Prinz, 2002; Thompson & Russell, 2004; Want & Harris, 2001; Whiten, Flynn, Brown, & Lee, 2006). From this research, it became clear that observing another's behaviour can be as effective, and often more effective for learning, than attempting a new task alone. However, this body of work has been mainly restricted to the dyadic context of a single observer and a single model. Culture is bigger than this. In particular, culture requires multiple, serial episodes of observational learning across a number of individuals. It is this phenomenon that diffusion studies specifically target.

In the present study, a diffusion chain design is used that compares the transmission of information across chains of five individuals when the first child in each chain is trained in one of two techniques. Such a design offers a controlled and parsimonious representation of culture, here used to examine age and gender effects in cultural transmission. We note that the 'diffusion experiment' literature offers alternative designs that include groups of models demonstrating different techniques, either to individual observers or to groups of observers, and in some contexts this may be a more ecologically realistic paradigm. However, there are also many contexts in which opportunities for transmission occur at the dyadic level, such as between parent and child or between two peers. This is the focus of the present study.

Our focus on peer-to-peer transmission is significant in itself. Many models of child psychology emphasize an expert-to-novice form of transmission, most notably Vygotsky (1981). In the present study, each demonstrator becomes an expert through observation of, and practice with, the task, yet these experts are peers rather than the commonly used adult-expert. Such a change in authority adds an interesting dimension to observational learning, as cultural learning through peer-to-peer transmission may be more relevant for some information than learning from adults. Learning from an expert peer has been of interest for many years (for example see Wood, Wood, Ainsworth, & O'Malley, 1995), and there is ample research on peer interactions and the role of peer influence (Hartup, 1996), which reinforces the importance of investigating this social relationship.

Diffusion studies have been extremely informative, but there is still much scope for improving their design. Some studies have omitted control groups, in which a target object is available but no model is present to demonstrate how the object can be manipulated (for example, Huffman & Hirata, 2004). Such controls are important in allowing an examination of both individual learning and the initial propensity to adopt certain behaviours rather than others. A critical strength of the current study is the inclusion of three groups: a no-model control condition and two experimental conditions in which alternative actions are introduced. This 'two-actions' aspect of the design (Dawson & Foss, 1965; Whiten, Custance, Gomez, Teixidor, & Bard, 1996) offers additional information about which learning mechanism underlies any cultural transmission documented. In a two-action design, initial models are trained to use one of two different techniques to achieve the same outcome. That both methods achieve the same result through different techniques allows a distinction to be drawn between emulative learning (copying results) and imitative learning (copying means) (see Want & Harris, 2002 for an extended discussion), as well as controlling for low-level observational learning effects like stimulus enhancement, in which the model merely draws the attention of the observer to the objects. Different cultural traditions can be said to exist if each of the techniques is transmitted along the respective chain, and if neither technique occurs in the no-model control condition, we can infer that significant cultural learning has taken place in the experimental groups. In a complementary way, diffusion experiments allow an examination of the cumulative effect of errors or innovations, thus illustrating how behaviour, and potentially traditions, may evolve over time (Mesoudi, Whiten, & Laland, 2006).

The present study aims to extend our understanding further by examining cultural transmission within the realm of tool use. Children's tool use *per se* has received much attention and a growing corpus of research examines the development of tool use to solve problems, from obtaining an out-of-reach toy by one-and-a-half- to four-year-olds (Chen & Siegler, 2000; Gredlein & Bjorklund, 2005; Want & Harris, 2001) to solving an insight problem in which an artefact needs to be used in an atypical manner, with five- to seven-year-old children (Defeyter & German, 2003). However, as Want and Harris (2002) have noted, little research has specifically investigated 'how children might learn to use tools from watching other people . . . . Developmental psychology . . . has almost completely neglected the social learning of tool-use (which typically involves learning the use of one object in conjunction with another) in human children. Instead developmental studies have focused mainly on the imitation of facial and manual gestures (most famously, tongue protrusion) or simple actions on objects (such as pushing a button on a box)' (pp. 1–2). The present experiment thus fills a void in the field of observational learning by (1) examining the transmission of information across groups of children; (2) focusing on child-to-child transmission rather than on experimenter-child dyads; (3) using a powerful three-group (two-action plus no-model control) design; and (4) focusing on the realm of tool use.

A small literature on observational learning of tool use exists, although limited to one-to-one transmission. Nagell, Olguin, and Tomasello (1993) studied children witnessing the raking-in of a previously out-of-reach, desirable object. In one demonstration the rake, which was resting on its prongs, was 'flipped' so that the complete edge was used for the object retrieval. In another demonstration, the rake was not flipped but was already resting on its complete edge. Irrespective of which demonstration a child witnessed, they were presented with a rake resting on its prongs. Children in a no-model control condition simply used the rake without flipping it over. Children from the flip condition initially did the same, but then flipped the rake over to use the more efficient, complete edge. Children who did not witness the rake being flipped over but saw a complete edge demonstration did not flip the rake over and used the inefficient prong edge. These findings were taken as evidence that children are relatively 'blind imitators' regardless of the utility of the technique. Similarly, Whiten et al. (1996), Horner and Whiten (2005), and McGuigan, Whiten, Flynn, and Horner (2007) found two- to four-year-old children imitated the exact form of the actions they witnessed and even copied irrelevant, that is, not functionally important, actions. All of these studies confirmed the feasibility of studying social learning of tool use in children and gained instructive results, but none examined the extent to which observational learning is sufficient to maintain traditions across multiple cultural transmission events.

In addition, we noted the relevance of a recent study examining how chimpanzees in the wild learn to use tools to fish for termites. Lonsdorf, Pusey, and Eberly (2004) found that males learned through direct experience in playing at a termite mound whereas females spent more time than males watching their mothers fishing for termites and successfully fished for termites at a younger age than males. The present study offers a unique opportunity to examine whether such gender differences seen in the learning style of our closest living relative, the chimpanzee, also exist in children. Gender differences related to learning reported in the child development literature (Blote & Van Gool, 1989; Blote & Van Haasteren, 1989; Vederhus & Krekling, 1996; Warrick & Naglieri, 1993) concern success at a task or cognitive processing rather than differences in learning style, as reported by Lonsdorf et al. (2004). To our knowledge, studies of gender differences in children's social learning have not reached a clear conclusion (Younger et al., 2005), although the general opinion is that girls prefer collaborative learning (Askew & Ross, 1988; Fennema, 1996) whereas boys prefer a more 'hands-on' learning experience (Barker, 1997). In the present study, investigating potential gender differences in the performance of children in both a no-model control and diffusion chain study provides an important opportunity for comparison with Lonsdorf et al.'s findings. If Lonsdorf et al.'s (2004) findings were replicated in children, the no-model control condition will see boys being more successful than girls, consistent with a 'hands on, trial and error' learning style whereas girls will be more successful than boys in the diffusion chains, where they have the opportunity to observe a demonstrator.

Accordingly, the task used in this study was designed to display certain analogies to termite fishing in so far as skilled use of a stick tool is required to extract small items from a protected space, and we called it the extractive tool-use task (ETT, Figure 1). We recognize the significant differences between termite fishing seen in wild chimpanzees and the task presented to the children in this study. Real termite fishing entails skilful insertion of a flexible probe into an opaque container, and the subsequent withdrawal,

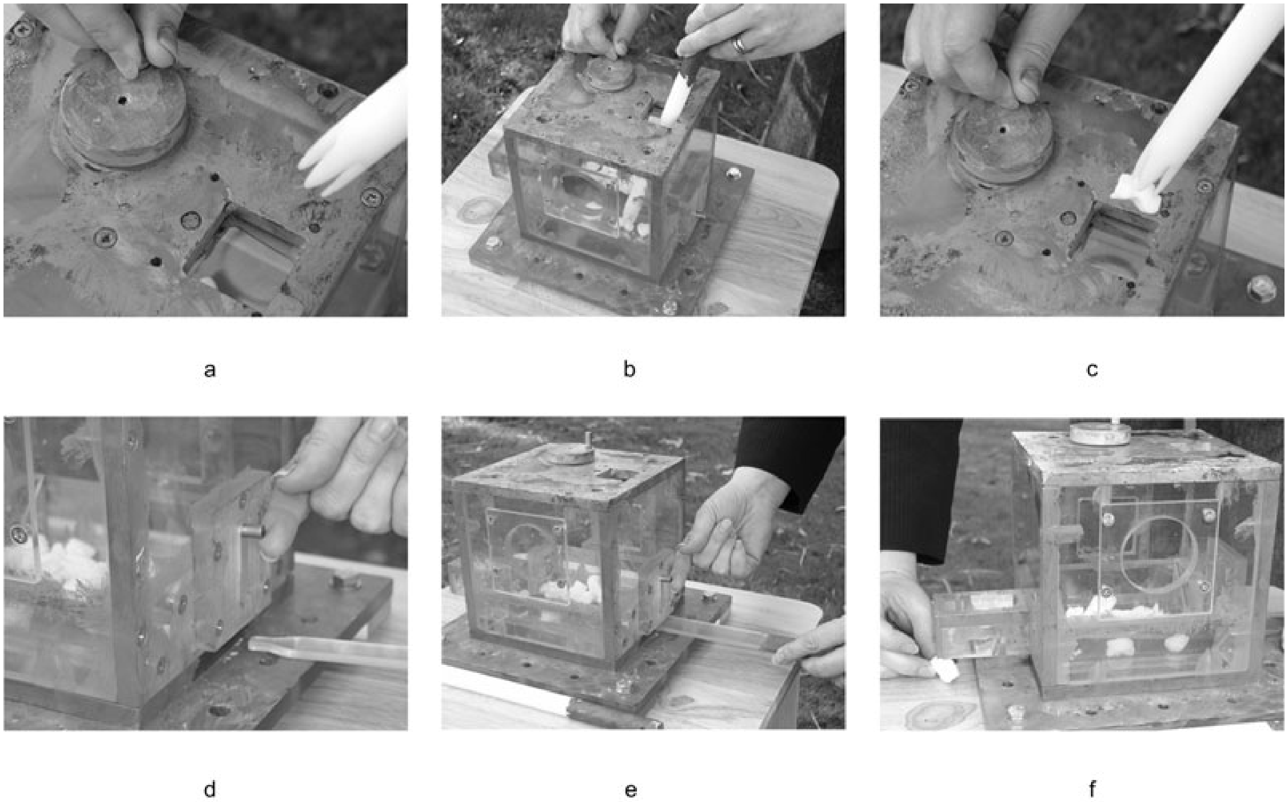

Figure 1. Alternative technologies transmitted in diffusion chains. (a–c), ‘Stab’ technique; (a), rotating knob to open window; (b), inserting tool stab polystyrene beads; (c), successfully withdrawing tool. (d–f), ‘Slide’ technique; (d), lifting hatch door; (e), inserting tool to push beads; (f), pushing beads out of chute.

all of which is based on tactile feedback. The ETT involves a transparent box containing polystyrene beads that can be removed using one of two techniques, 'stab' vs. 'slide'. Both techniques use a comparable complexity of actions including taking hold of a tool, opening a small door, inserting a tool into the ETT, and using the tool to remove a bead, and they result in the same outcome, removing a bead.

As well as examining gender differences in tool use, the present study investigated developmental changes in children's individual and observational learning. It might be predicted that with their superior memory skills, older children (five-year-olds in the present study) are better at copying what they witness than younger children (three-year-olds). Alternatively, five-year-olds may use these additional memory skills to free cognitive resources to allow more innovation on the task whereas three-year-olds may be more dependent on, and therefore faithful to, the demonstration they have witnessed. Thus, we made no predictions about whether we would discover greater fidelity of transmission in the younger or older children.

Our core research question addressed whether, when employing an appropriate task for this age group, fidelity of transmission would be sufficient to sustain alternative traditions across multiple 'cultural generations' in the chains and what kinds of deviation from this might occur. Further, how is this transmission related to age and gender?

Method

Participants

One hundred and twenty-seven children participated, 65 three-year-olds and 62 five-year-olds. Of these, 47 children (25 three-year-olds and 22 five-year-olds) were allocated to a no-model control condition and 80 children (40 three-year-olds and 40 five-year-olds) were allocated to diffusion chains. Each chain contained children of the same age group and gender. The latter replicated a tendency in young children to play in same-sex groupings (Bishop & Curtis, 2001) and facilitated comparison of gender differences in transmission. The sequence of children in each chain ran from eldest to youngest. The mean age of the three-year-old children in each chain ranged from 3 years and 8 months to 3 years and 11 months (standard deviation [*SD*] ranged from 1.5–4 months), and the mean age of the 5-year-old children in each chain ranged from 5 years and 7 months to 5 years and 10 months (*SD* ranged from 1.5 months to 3 months). The mean age of the three-year-old children in the no-model control condition was three years and eight months ($SD = 2$ months) and the mean age of the five-year-old children in the no-model control condition was five years and eight months ($SD = 2$ months).

Design

The study used a between-group, diffusion chain design to compare observational and individual learning in relation to age (three- vs. five-year-olds) and gender. Children were allocated to a no-model control condition or a diffusion chain, each chain containing five children in total. Chains differed according to age (three- vs. five-year-olds), gender, and the technique witnessed (stab vs. slide, explained below) to obtain a polystyrene bead from the ETT. Two chains were run for each of the eight conditions defined by these three factors, yielding 16 chains in all.

Materials

The ETT was a transparent cube approximately 16 cm on each side. It was bolted to a table so that it could not be moved. Inside the ETT were polystyrene beads, often used as packaging material. On the top of the ETT was a handle that if turned clockwise opened a 3 cm-wide square door next to it (see Figure 1a). The door was spring-loaded so that when the handle was released the door would automatically close. On one side of the ETT was a second handle, which lifted a second door revealing a horizontal slot just above the bottom of the ETT (see Figure 1d). This mechanism was not spring-loaded, but if the handle were released, the door would close due to gravity. Opposite the slot was a chute that protruded away from the box and did not contain a bottom panel so that anything pushed through the chute would fall onto the top of the table. The only other opening was a door through which the experimenter could re-bait the ETT and which was kept firmly locked while participants took part in the study. As well as the ETT, two tools were placed on the table. One tool was a plastic rod (27 cm in length), one end of which was cut so that it had four prongs, and the other end was cut flat and painted brown. This 'stab' tool was designed to be put into the top door so that a polystyrene bead could be stabbed and then removed through the top door (see Figure 1a–c). The second tool was a long, thin, ruler-like tool (30 cm in length) with a tapered end, with the other end painted brown. This 'slide' tool was designed to be placed through the horizontal slot, so the tapered end could guide a polystyrene bead out of the chute (see Figure 1d–f). Both tools had only one end that was functional, that is, only one end of the stab tool had prongs and only one end of the slide tool could fit through the chute.

Procedure

Each participant took part in two separate test sessions. During the first session, the children were tested using the British picture vocabulary scale (Dunn, Dunn, Whetton, & Pintilie, 1997) to check that those included in the study were in the typically developing range. No child needed to be removed from the study.

During the second session, children took part in either the diffusion chains or no-model control condition. Children were allocated to a specific position in each chain of five children before a testing session began. Each chain contained five children of the same age group (three- vs. five-year-olds) and gender. Chains were matched so that children of similar ages appeared at the same position in each chain. Testing took place in a quiet room away from the other children in the school or nursery. Initially, the experimenter said to the first child, 'okay watch me and then you can have a go'. Then the child watched the experimenter retrieve a polystyrene bead from the ETT using one of two techniques, stab or slide. Having witnessed two demonstrations, the child was allowed to have a turn, 'now it's your turn'; the goal of retrieving a polystyrene bead was never explicitly stated. For all attempts, both tools were placed on the table next to the box; they were never physically given to the child. A turn was said to end when a tool was removed from the ETT, whether a bead had been removed or not. The first child in every chain attempted to retrieve a bead on both of their two attempts and all attempted to use the technique they had been shown. After the first child's two attempts, the second child in the chain was brought into the room and told to wait while the first child had two attempts, then it would be his/her turn. No explicit instructions were given about watching or teaching, but the experimenter made sure that each child had a clear view of the ETT and the actions upon it. After the first

Table 1. The Achievement Score Criteria for all Children's Performance on the ETT

Score	Behaviour
0	Touched nothing
1	Touched either tools or doors, but not both
2	Touched tools and doors
3	Opened door, and also touched tools, but never inserted tools
4	Inserted tool but did not touch beads
5	Touched beads but not functionally
6	Moved a bead functionally with the tools but did not successfully remove it
7	Success: at least one bead removed

ETT = extractive tool-use task.

child's two demonstrations, s/he returned to her/his classroom and the child, who had been present during the demonstrations, had two solo attempts before becoming a demonstrator for the next child in the chain. This procedure continued to the fifth, and final child, who had only two attempts as there was no need for them to demonstrate.

Each child's performance was scored on a scale of 0–7 for the two alternative techniques (see Table 1). In some cases, children produced all the appropriate actions to succeed but still failed to retrieve the bead. For example, a child stabbed a bead with the stab tool, only to have the bead fall from the end of the tool as it was being taken out of the top door. Such a child (scoring 6) was deemed to be an appropriate model for the next child in the chain. By contrast, children who inserted a tool into a hole but did not touch the polystyrene beads or touched the beads but not functionally were not recruited as a model, as they were not deemed to have mastered the task sufficiently. This specific form of behaviour (awarded achievement scores of 4 or 5, see Table 1) occurred on less than 2 percent of all the attempts. Children were dropped from a chain if they had an achievement score of 5 or below, in which case the previous child in the chain was brought back and demonstrated for the following child, that is, if the third child dropped out, then the second child would return to demonstrate to the fourth child. It should be emphasized that a child was allowed to be a model irrespective of the technique s/he used as long as they achieved a score of 6 or 7. Therefore, a child who used the stab technique in her/his attempts, even though she had witnessed the slide technique, could be a model for the next child in the chain. All children who acted as models more than once achieved the highest level during their first demonstrations (obtaining an achievement score of 7) and in all their subsequent demonstrations. Therefore, the model's performance was not overtly affected by practice.

In the no-model control condition, children were brought into the room and presented with the ETT and tools, being told, 'lots of boys and girls have had a go, and now it's your turn'. Testing ended if a child successfully retrieved a bead, refused to continue after general encouragement, or after 12 minutes of interaction with the ETT. Children who struggled in the diffusion chains and no-model control condition were given general encouragement, including, 'what do you think you do now?', 'you can touch it as much as you like, you can't break it', and 'you're doing really well, what do you think you do next?'.

All children, irrespective of success, received a sticker as a reward at the end of the testing session. Although children across both the diffusion chains and the no-model control condition were never explicitly told to retrieve a bead from the ETT, results from previous work (Horner et al., 2006), where the overall goal was never made explicit, show children are motivated to participate and able to infer the aim of the task.

All the testing sessions were videotaped. Coding of the videotapes discriminated the technique used to retrieve a bead, the tools and doors touched, the level of success, and the time taken at each attempt or demonstration. The level of behaviour a child achieved was divided further by the distinctions shown in Table 1. An independent observer, who was blind to the rationale of the study, coded 22 percent of the sample (28 children). All Kappa scores (tools touched, doors touched, technique used, attempt success, demonstration success, and removal of child) were .85[1] or above, showing a good level of reliability.

Results

We first describe the behaviour of children in the no-model control condition and examine the role of age and gender in individual learning to which this condition was constrained. Comparing the performance of the children in the no-model control condition with those in the diffusion chains then allows conclusions to be drawn about individual vs. observational learning. The performance of the children in the diffusion chains is then presented in more detail, discriminating age and gender effects. Changes in behaviour across the 'cultural generations' of the chains are then investigated. Finally, we examine the effect of the level of success of the demonstrations witnessed on a child's own attempts. In order to give an overview of the data, Table 2 presents a summary of the children's performances broken down by condition, age, gender, and, where appropriate, technique witnessed. It should be noted that 12 children failed to become appropriate models for the following child in the chain, as they performed no actions that were purposeful. Therefore, some of the data analyses for the diffusion chains do not contain the full 80 children, and in such cases, the number of children who were included is explicitly stated.

No-model Control

A binary logistic regression found no significant main effects for age ($B = .89$, $SE = .92$, $Wald = .95$, NS) or gender ($B = 1.49$, $SE = 1.98$, $Wald = .57$, NS) on the no-model control children's level of success, and also no significant interaction between age and gender on success ($B = -.38$, $SE = 1.25$, $Wald = .09$, NS), although there were trends towards five-year-olds (success rate = 50 percent) being more successful than three-year-olds (32 percent), and boys (52 percent) being more successful than girls (29 percent). Similarly, there were no significant gender (F [1, 47] = .61, NS) or age (F [1, 47] = 3.51, NS) effects in the children's achievement scores, nor any interaction (F [1, 47] = .44, NS).

Fifteen of the 19 control children who were successful used the stab technique. The other four children (five-year-olds, three boys and one girl) put the slide tool in the top hole and then successfully used it to flick a bead out of the chute. No child used the slide technique. Children were given as much time as they wanted to play with the ETT. Five-year-olds ($M = 4$ minutes 26 seconds) spent approximately 50 percent

Table 2. Mean Achievement Scores and Rates of Performance

	Three-year-olds				Five-year-olds			
	Female		Male		Female		Male	
Diffusion Chains	Witnessed stab	Witnessed slide	Witnessed stab	Witnessed slide	Witnessed stab	Witnessed slide	Witnessed stab	Witnessed slide
Mean achievement score								
Attempt 1	5.80 (10)	5.70 (10)	6.20 (10)	5.90 (10)	6.20 (10)	4.70 (10)	6.80 (10)	6.90 (10)
	1.62	*1.49*	*.63*	*1.52*	*1.55*	*2.26*	*.42*	*.32*
Attempt 2	6.38 (8)	6.50 (8)	6.50 (10)	6.63 (8)	6.89 (9)	7.00 (5)	6.80 (10)	6.90 (10)
	.74	*.53*	*.53*	*.52*	*.33*	*0*	*.42*	*.32*
Demonstration 1	6.33 (6)	6.57 (7)	7.00 (8)	6.86 (7)	6.86 (7)	7.00 (5)	7.00 (8)	7.00 (8)
	.82	*.53*	*0*	*.38*	*.38*	*0*	*0*	*0*
Demonstration 2	6.33 (6)	6.71 (7)	6.75 (8)	6.71 (7)	6.86 (7)	7.00 (5)	6.75 (8)	7.00 (8)
	.82	*.49*	*.46*	*.49*	*.38*	*0*	*.46*	*0*
% success rate								
Attempt 1	40	30	30	40	60	40	80	90
Attempt 2	50	50	50	63	89	100	80	90
Demo 1	50	57	100	86	86	100	100	100
Demo 2	50	71	75	71	86	100	75	100
No-model control	Three-year-olds				Five-year-olds			
	Female		Male		Female		Male	
Mean achievement scores	3.14 (14)		5.00 (11)		4.20 (10)		5.08 (12)	
% success rate	21		45		40		58	

Note: Numbers in parenthesis refer to the number of participants in each group, and the numbers in italics refer to the *SD*.

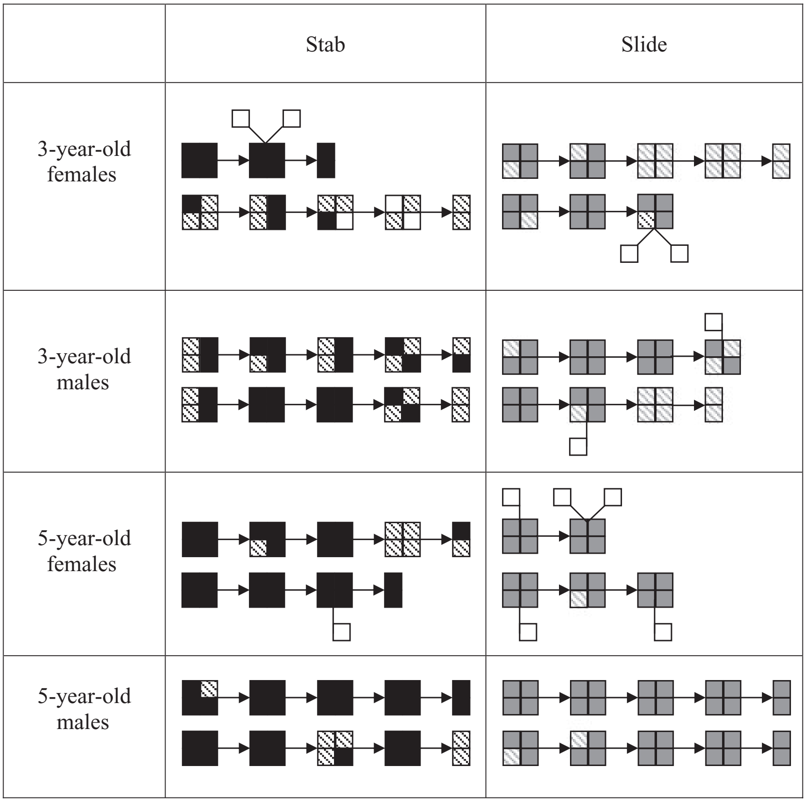

Figure 2. Transmission along diffusion chains: each set of four squares represents the first two attempts (left-hand side) and the two demonstration trials (right-hand side) for one child; black represents successful stab technique (achievement score of 7), grey represents successful slide technique (achievement score of 7); hatched squares represent achievement scores of 6 for either technique; white represents achievement score of 5 or less. Children who remained in the chains but had trials represented by a white square, attempted to remove the bead in another trial; children who dropped out from the chains and appear as an offshoot of the chain did not attempt to remove a bead.

longer on the task than the three-year-olds ($M = 2$ minutes 58 seconds; F (1, 47) = 6.49, $p < .05$), and girls spent on average 33 percent longer on the task than boys (girls' $M = 4$ minutes 15 seconds, boys' $M = 3$ minutes 2 seconds; F (1, 47) = 5.02, $p < .05$). There was no significant interaction between age and gender in the time taken (F [1, 47] = .39, NS).

Diffusion Chains

Figure 2 represents the behaviour of the children in the diffusion chains. Comparisons of children seeded with the stab or slide technique found no significant difference in the ability to retrieve a bead (attempt 1, $\chi^2[1, 80] = .50$; attempt 2, $\chi^2[1, 68] = .36$; demonstration 1, $\chi^2[1, 56] = .02$; demonstration 2, $\chi^2[1, 56] = 1.36$), the likelihood of dropping out of the chain due to lack of ability ($\chi^2[1, 80] = 3.53$) nor achievement scores (averaged over all attempts[2], z[80] = .56). Because of these similarities, data were collapsed across the two techniques for the following comparisons of the no-model control condition and the diffusion chains. The only difference was that children in the slide condition (demonstration time $M = 49$ seconds) took longer to demonstrate than the children in the stab condition (demonstration time $M = 31$ seconds, F [1, 57] = 11.16, $p < .01$).

Observational Learning (1): Diffusion Chains vs. No-model Controls. Children in the diffusion chains were found to be significantly more successful at their initial attempt (success rate = 73 percent) than children in the no-model control condition (40 percent, $B = -1.51$, $SE = .42$, $Wald = 12.71$, $df = 1$, 127, $p < .001$). There was no significant main effect nor interaction for age and/or gender when rate of success of all 127 children was analysed. Comparisons of all children showed that children in the diffusion chains ($M = 6.03$, median = 7.00) had significantly higher achievement scores than children in the no-model control condition ($M = 4.29$, median = 5.00; F [1, 127] = 13.97, $p < .001$); again there were no significant interactions with age or gender. Comparisons of the children's manipulations of the ETT's doors and tools revealed different behaviours for children in the diffusion chains and those in the no-model control condition. Children in the diffusion chains tended to touch only one door; no-model controls were more likely to touch both doors ($\chi^2[4] = 36.25$, $p < .001$). Only one child, who was in the no-model control condition, touched neither door. These effects were found for both age groups (three-year-olds, $\chi^2[4] = 12.09$, $p < .05$; five-year-olds, $\chi^2[3] = 29.46$, $p < .001$) and gender groups (girls, $\chi^2[4] = 20.07$, $p < .001$; boys, $\chi^2[2] = 16.93$, $p < .001$). Similarly, children in the diffusion chains were more likely than the no-model controls to touch only one tool ($\chi^2[2] = 35.81$, $p < .001$). Children in the no-model control condition were also more likely to touch neither tool ($\chi^2[3] = 43.02$, $p < .001$). These effects were found for both age groups (three-year-olds, $\chi^2[3] = 21.60$, $p < .001$; five-year-olds, $\chi^2[3] = 23.85$, $p < .001$) and gender groups (girls, $\chi^2[3] = 22.58$, $p < .001$; boys, $\chi^2[3] = 21.63$, $p < .001$).

Children in the diffusion chains were also significantly quicker in their solo attempts (attempt time $M = 1$ minute 3 seconds, F [1, 127] = 97.02, $p < .001$) and in their demonstrations (demonstration time $M = 40$ seconds, F [1, 104] = 99.99, $p < .001$) than children in the no-model control condition (attempt time $M = 3$ minutes 39 seconds). The children in the no-model control condition who were successful at the task spent significantly less time on the task ($M = 2$ minutes 48 seconds) than children in the no-model control condition who were not successful ($M = 4$ minutes 14 seconds, F [1, 47] = 4.67, $p < .05$). Contrasts of the diffusion chain children's attempt and demonstration times showed that children were significantly faster in their later demonstrations than in their initial attempts (F [1, 56] = 15.27, $p < .001$).

Observational Learning (2): Comparing Groups Seeded with Different Techniques. All the successful children modelled the technique they had witnessed (see Figure 2). There was no variation in overall technique, apart from one child (three-year-old, female) who had witnessed the slide demonstration and who tried to complete a stab during her first attempt (she failed) but then demonstrated slide to the next child (see Table 3; $\chi^2[1] = 68$, $p < .001$).

Observational Learning (3): Age and Gender Effects. Twelve children dropped out from the chains, as they did not meet the criterion to act as a model for the next child

Table 3. The Number of Children Who Used the Same/Different Technique to That Witnessed

	Attempted only witnessed technique	Attempted other technique
Witnessed stab	37	0
Witnessed slide	30	1

in the chain. These children were four three-year-old females, two three-year-old males, and six five-year-old females. There was no age effect in the likelihood of dropping out ($\chi^2[1] < .001$); however, there was a gender difference, with more girls dropping out than boys ($\chi^2[1] = 6.26$, $p < .05$, see Figure 2). The achievement scores of the 12 children who dropped out ranged from 2 to 5, with a mean score of 2.83. The children who dropped out typically touched both the tools and the doors, but did not open the doors, and rarely inserted a tool.

For the 68 children who remained in the chains, scores could be given for both their first and second attempts, and for all those not at the end of a chain, for their first and second demonstrations. The achievement scores were compared across these four attempts for age and gender effects. Although average scores over all four attempts or over attempts and demonstrations could be used, it was felt that performance may change over the four attempts and so they should be analysed separately. Age was found to have a significant effect across all attempts and demonstrations (attempt 1, $z[68] = -3.23$, $p < .01$; attempt 2, $z[68] = -3.19$, $p < .01$; demonstration 1, $z[56] = -2.28$, $p < .05$; demonstration 2, $z[56] = -1.97$, $p < .015$), with five-year-olds having significantly higher achievement scores than three-year-olds, as illustrated in Table 2. A significant gender effect was found in the first demonstration ($z[56] = -2.62$, $p < .01$), with boys having significantly higher achievement scores than girls. This gender effect disappeared in the second demonstration. Examining the age groups separately showed that the gender difference found in the first demonstration remained significant for three-year-olds ($z[28] = -2.38$, $p < .05$), but not for five-year-olds.

There were no age or gender differences in the average time taken to complete the first two attempts. There were no gender differences in the average time taken to complete the demonstrations, but five-year-olds ($M = 30$ seconds) were quicker than three-year-olds ($M = 49$ seconds; $F [1, 56] = 12.43$, $p < .001$).

Observational Learning (4): Change over Generations in the Chain. Spearman correlations were computed to identify any changes in achievement scores over generations along the chain. These were initially carried out across all the data, and then for separate age groups. Overall, the achievement scores were related to a child's position in the chain. Correlation with generation position approached significance for an averaged attempt score (Spearman's rho$[68] = -.23$, $p = .06$) and predicted an averaged demonstration score (Spearman's rho$[56] = -.34$, $p < .05$); therefore, although children showed great fidelity by using the same technique within individual chains, in later generations within the chains children were less likely to implement the method successfully.

Table 4. Relations between the Types of Demonstrations Children Witness and Their Own, Later Attempts

		Attempted fail only	One fail, one success	Success only
Witnessed	Fail only	6	1	0
	One fail, one success	1	1	3
	Success only	4	19	33

These associations could have been influenced by age, as chains ran from oldest to youngest. Therefore, the same analyses were undertaken for three-year-olds and five-year-olds separately, with and without age partialled out. For the five-year-olds, none of the above effects were apparent even before age was partialled out. For three-year-olds, generation position correlated significantly with the second attempt achievement scores (Spearman's rho[34] = –.34, $p < .05$) and the second demonstration achievement scores (Spearman's rho[28] = –.64, $p < .001$), but not the first attempt achievement score or the first demonstration achievement score. However, both of the significant effects disappeared when age was partialled out.

Observational Learning (5): Learning from Successes and Failures. Finally, we investigated whether there is an association between the type of attempt a child witnessed, that is, successful vs. unsuccessful, and a child's own, later attempts. Unsuccessful attempts occurred on 20 of the 80 witnessed attempts (25 percent). In this analysis it is appropriate to include only children who remained in the chains, as one cannot be sure that children who dropped out were actually watching the previous child's attempts. Of the 68 children included in the chains, there was a significant association between what children witnessed and what they attempted ($\chi^2[4] = 28.91$, $p < .001$), as shown in Table 4.

This association was explored further by using the achievement scores described in Table 1. A correlation between an average of the achievement a child had witnessed and an average of the achievement he or she produced was significant (Spearman's rho[68] = .37, $p < .005$) and remained significant when age was partialled out ($r[56] = .41$, $p < .001$).

Discussion

The primary goal of this study was to extend the application of the diffusion chain paradigm to young children and explore what it can tell us about cultural transmission in this population. Building on an extensive literature documenting imitation and other kinds of observational learning in young children studied at the dyadic (model and observer) level, we adopted the diffusion chain approach as one variant of a more realistic model of the larger picture of culture.

Our core questions were thus about whether, when employing an appropriate task, fidelity of transmission would be sufficient to sustain alternative traditions across multiple 'cultural generations' in the chains and what kinds of deviation from this might occur. Initially, it is important to consider if the task was indeed appropriate. The principal criteria for this must be that there were no floor or ceiling effects for age or gender. This was the case in so far as none of the subgroups (age and/or gender) displayed a total inability to complete the task, either in the no-model condition or the observational learning conditions; nor did any of the subgroups perform at ceiling level. In the control condition, 32 percent of three-year-olds and 50 percent of five-year-olds and 52 percent of boys and 29 percent of girls were able to achieve success in the task without benefit of a model. These figures might suggest that the task was, if anything, too easy for an experimental analysis of the additional effect of observational learning over baseline levels. However, given that several diffusion chains in our observational learning conditions displayed significant failures, even at the older age, the level of difficulty in the task appears well judged. In future studies, it will be of interest to incorporate tasks with different levels of difficulty in the no-model control condition, according to age and gender.

Whether there would be evidence of cultural transmission along chains was answered strongly in the affirmative and at several levels of analysis. The effect is most dramatic in comparing the two conditions defined by use of one technique rather than the other. There was in fact no cross-over between these, with the single exception of a three-year-old female, who had witnessed the slide demonstration, tried unsuccessfully to complete a stab during her first attempt but then demonstrated slide to the next child (see Table 3); all chains starting with stab ended with the same technique, as did those starting with slide. In this respect, all subgroups displayed an extreme level of cultural conformity to the approach they witnessed, without any encouragement from the experimenters to do so. It could be argued that for innovation to take place, a child needs more experience of the task than the present study allowed. Evidence against this is that a number of children in the no-model control condition did show innovation on their single attempt, using a 'flicking' method not anticipated by the experimenters, and discussed further below.

As this is the first child diffusion study of its kind, direct comparison within the developmental literature is not possible. However, the results are consistent with earlier dyadic studies that showed remarkable degrees of conformity, such as Meltzoff's demonstration of young children's willingness to copy even quite bizarre actions, such as pushing a buzzer switch with one's forehead (Meltzoff, 1995b), and Horner and Whiten's (2005) documentation of copying acts that are visibly unrelated to outcomes of interest. Only three other studies have employed our three-group, two-action diffusion design allowing some comparison across other species; two of these report lower levels of fidelity than we report here. Galef and Allen (1995) seeded two chains of rats with preferences for differently flavoured foods and found that these were sustained over several generations but were progressively degraded. Whiten et al. (2005) seeded two alternative kinds of tool use in an open diffusion study with chimpanzees and also documented differential spread, but some corruption to the opposite technique occurred in both groups. Horner et al. (2006) compared the transmission of information in children and chimpanzees using a puzzle box that did not require tool use, but in which a door could be lifted or slid open. Fidelity of the method used across generations was found to be high in both species, the child data being consistent with the results of the current study. Importantly, the present study extends this work by using a tool-use task and, specifically, by addressing gender and age effects.

A further source of graphic evidence of the power of cultural transmission among the children studied comes from comparisons between the diffusion chains and the no-model controls. Three effects deserve to be highlighted. Firstly, children in the experimental groups achieved significantly more success than those in the no-model control groups. Secondly and less predictably, one of the techniques, slide, was totally absent from the control groups. By contrast, it was the only technique employed by those children in chains seeded with slide. As noted above, it was also absent in the stab groups, and the three-way comparison between the slide, stab, and control groups thus demonstrates, dramatically, the strength of the capacity of cultural transmission in these young children. Thirdly, four control children invented a technique that involved inserting the slide tool through the top hole and flicking a bead into the chute and out of the apparatus. This was not seen in any of the observational learning groups, in which children simply adopted either slide or stab as witnessed. This illustrates the power of cultural transmission to 'canalize' behaviour patterns into narrower channels than might otherwise occur.

However, in contrast to the strength of all these kinds of evidence of high-fidelity transmission across multiple cultural generations, Figure 2 also shows some converse evidence, for although a majority of chains (nine) ran to the fifth generation, seven chains did not. This suggests some qualification to the discussion above, highlighting strong cultural transmission effects. Most failures appeared to represent not so much limits on observational learning, but on either the motivation or the competence to complete the task in the way observed. Most children who did not achieve the maximal score of 7 were performing sufficiently well (score 6) to act as models and all were attempting the technique witnessed. Of the 12 who achieved less, 5 could be assigned to one technique, as they touched only the tool and door associated with the technique seeded in their chain. The other seven behaved like no-model control children and touched both doors and/or both tools. For these children, it remains unclear whether their lack of success was due to personal limits in factors such as motivation or competence, or was instead due to limits on observational learning (perhaps they had not noticed key aspects of the technique they had watched). The present study was not designed to discriminate such factors and additional experiments will be required in future to do so.

Our finding that a child's level of achievement was correlated with that of the child s/he had previously watched does pick out an observational learning effect. However, the reason why a model is relatively poor in the first place might reflect either his/her own prior observational learning experience (perhaps s/he had watched a poor model) or personal factors like competence and motivation, as discussed above.

The finding that transmission is influenced by the quality of the modelling observed might be seen as in conflict with some earlier studies of observational learning that emphasized the capacity of children to learn from sub-standard models. Notably, Meltzoff (1995a) showed that infants successfully completed an action that they had seen a model fail at during the model's attempt. For example, a model clumsily draped beads over the side of a cup, yet the infant later placed the beads more neatly in the cup. Want and Harris (2001) showed that young children would learn better from a demonstration that illustrated mistaken ways to proceed than successful demonstrations alone, although in this case the incorrect approach was immediately followed by a correct one. However, there is not necessarily a conflict with our results. An ability to learn from failed attempts does not mean that better performance will not follow from watching successful attempts. In any case, although there was an association between a model's success and the observer's success, cases were recorded where failed attempts were sufficient for transmission of techniques to continue. For example, in one chain of three-year-old girls, the slide technique was transmitted from the third to the fifth generation with no intervening success, and in one chain of five-year-old girls, success was achieved using the stab technique in the final generation despite a lack of success in the prior generation.

In the present study, we were interested in the transmission, and in failures in transmission, of techniques across chains of children. There is evidence that children from the middle of their second year will copy the specific means an adult uses to produce a particular outcome, even if a more efficient method is available (Call, Carpenter, & Tomasello, 2005; Horner & Whiten, 2005; Nielsen, 2006). Our results support such findings, for although the two techniques were roughly comparable in complexity, requiring similar numbers of steps, only the stab technique appeared in the no-model control group. Future work could explore the transmission of information across chains, when one of the techniques is overtly more complex than an alternative technique, seeded in a separate chain.

More specific goals of our study were to examine the possibility of gender and age differences in cultural transmission in three- to five-year-olds. Unexpected effects were discovered and are best discussed together. If age is considered independently, the differences were not dramatic. There was in fact no significant age difference in the likelihood of children dropping out of chains. However, for those who remained in the chains, five-year-olds had slightly but significantly higher scores than three-year-olds, which suggests that this effect relates to the capacity for observational learning in so far as an age effect was not significant in the achievement scores of the no-model control children. The age difference in children's observational learning capacity may be specific to this ability; alternatively, it may reflect differences in other cognitive skills such as inferring mental states to other's actions, which undergoes significant changes between the ages of three and five years. Our current research is addressing this issue.

The gender differences identified were more surprising. Recall that one major reason for investigating this issue was the demonstration that young female chimpanzees show more commitment to observational learning of tool use and master it earlier (Lonsdorf et al., 2004). In contrast, in our study it was males who achieved higher levels of success. Girls were more likely to drop out of chains, and those who remained in the chains had lower achievement scores. However, it is important to note that the difference in achievement score is true of only one of the children's four individual attempts at the task—their first demonstration. Such an effect may be due to motivational factors as there is no deficit in girls' competence on earlier attempts at the task compared to boys. Perhaps girls are more reticent in their first social interaction with the next observer child. However, this cannot explain the higher dropout for girls, which occurred through their low solo test achievement, most noticeable in five-year-old females who witnessed the slide method (Figure 2). Further research on these issues is required. One obvious future step would be to discover if other kinds of activities than tool use reverse the gender difference in sustainability of successive social transmissions that we found in the present study.

In conclusion, the current study has shown that it is possible to use diffusion studies with young children to investigate cultural transmission. Both three- and five-year-olds are capable of observing a specific technique within a tool-use task and of adopting the technique they have witnessed, transmitting this across generations. Five-year-olds are more proficient than three-year-olds at transmitting the technique faithfully across multiple generations, as three-year-olds produce more unsuccessful attempts. Yet neither three- nor five-year-olds attempt to innovate when they have observed another child complete the task using a particular technique, even when the witnessed technique is one that they would not adopt when presented with the task in an individual learning scenario and when children work alone on the task they are capable of innovation. Girls show poorer transmission than boys as they show worse performance when initially attempting the task in front of a naïve peer. Girls are also less likely than boys to complete the task, even though they have been present and had a clear view of a familiar peer having two attempts at completing the task.

The current study presents a first step towards ascertaining the effects of peers on cultural transmission by using same-age and same-sex groupings. To further this area of investigation, it will be necessary to compare the dynamic interactions among children in mixed-age and gender groups and examine additional factors such as perceived knowledge, social status and popularity, and friendships. Such an investigation will allow an in-depth analysis of peer-to-peer transmission of information in a naturalistic setting, addressing issues in observational learning that have been of interest in tutoring and collaborative scenarios.

References

Askew, S., & Ross, C. (1988). *Boys don't cry: Boys and sexism in education*. Milton Keynes: Open University Press.

Bandura, A., Ross, D., & Ross, S. A. (1961). Transmission of aggression through imitation of aggressive models. *Journal of Abnormal and Social Psychology*, *63*, 575–582.

Bangerter, A. (2000). Transformation between scientific and social representations of conception: The method of serial production. *British Journal of Social Psychology*, *39*, 521–535.

Barker, B. (1997). Girls' world or anxious times: what's really happening at school in the gender war? *Educational Review*, *49*, 221–228.

Bartlett, F. C. (1932). *Remembering*. Oxford: Macmillan.

Bishop, J. C., & Curtis, M. (Eds.). (2001). *Play Today in the Primary School Playground*. Buckingham: Open University Press.

Blote, A. W., & Van Gool, H. (1989). Writing behavior of children aged 4 to 5 and a half years. *Journal of Human Movement Studies*, *17*, 133–152.

Blote, A. W., & Van Haasteren, R. (1989). Development dimensions in the drawing behavior of pre-school children. *Journal of Human Movement Studies*, *17*, 187–205.

Boyd, R., & Richerson, P. (1985). *Culture and the Evolutionary Process*. Chicago, IL: University of Chicago Press.

Call, J., Carpenter, M., & Tomasello, M. (2005). Copying results and copying actions in the process of social learning: Chimpanzees (Pan troglodytes) and human children (Homo sapiens). *Animal Cognition*, *8*, 151–163.

Chen, Z., & Siegler, R. S. (2000). Across the great divide: Bridging the gap between understanding of toddlers' and older children's thinking. *Monographs of the Society for Research in Child Development*, *65*.

Curio, E., Ulrich, E., & Vieth, W. (1978). Cultural transmission of enemy recognition: One function of avian mobbing. *Science*, *202*, 899–901.

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

Defeyter, M. A., & German, T. P. (2003). Acquiring an understanding of design: Evidence from children's insight problem solving. *Cognition*, *89*, 133–155.

Dunn, L., Dunn, L., Whetton, C., & Pintilie, D. (1997). *British picture vocabulary scale*. Windsor: NFER-Nelson.

Fennema, E. (1996). Scholarship, gender and mathematics. In P. Murphy, & C. Gipps (Eds.), *Equity in the classroom: Towards effective pedagogy for girls and boys* (pp. 73–80). London: Falmer Press.

Galef, B. G., Jr., & Allen, C. (1995). A new model system for studying animal tradition. *Animal Behaviour*, *50*, 705–717.

Gergely, G., Bekkering, H., & Kiraly, I. (2001). Rational imitation in preverbal infants. *Nature*, *415*, 755.

Gredlein, J. M., & Bjorklund, D. F. (2005). Sex differences in young children's use of tools in a problem-solving task: The role of object-oriented play. *Human Nature*, *16*, 211–232.

Hartup, W. W. (1996). The company they keep: Friendships and their developmental significance. *Child Development*, *67*, 1–13.

Horner, V., & Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (*Pan troglodytes*) and children (*Homo sapiens*). *Animal Cognition*, *8*, 164–181.

Horner, V., Whiten, A., Flynn, E., & de Waal, F. B. M. (2006). Faithful replication of foraging techniques along cultural transmission chains by chimpanzees and children. *Proceedings of the National Academy of Sciences*, *103*, 13878–13883.

Huang, C. T. C., & Charman, T. (2005). Gradations of emulation learning in infants: Imitation of action on objects. *Journal of Experimental Child Psychology*, *92*, 276–302.

Huang, C. T., Heyes, C., & Charman, T. (2002). Infants' behavioral reenactment of 'failed attempts': Exploring multiple organ failure. *Developmental Psychology*, *38*, 840–855.

Huffman, M. A., & Hirata, S. (2004). An experimental study of leaf swallowing in captive chimpanzees: Insights into the origin of a self-medicative behavior and the role of social learning. *Primates*, *45*, 113–118.

Hurley, S., & Chater, N. (2005). *Perspectives on imitation: From mirror neurons to memes*. Cambridge, MA: MIT Press.

Laland, K. N., & Plotkin, H. C. (1990). Social learning and social transmission of foraging information in Norway rats (*Rattus noregicus*). *Animal Learning and Behavior*, *18*, 246–251.

Lefebvre, L. (1986). Cultural diffusion of a novel food-finding behaviour in urban pigeons: And experimental field test. *Ethology*, *71*, 295–304.

Lonsdorf, E. V., Pusey, E. A., & Eberly, L. (2004). Sex differences in learning in chimpanzees. *Nature*, *428*, 715–716.

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

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

Meltzoff, A. N. (1995b). What infant memory tells us about infantile amnesia: Long-term recall and deferred imitation. *Journal of Experimental Child Psychology*, *59*, 497–515.

Meltzoff, A. N., & Prinz, W. (2002). *The imitative mind: Development, evolution, and brain bases*. Cambridge: Cambridge University Press.

Menzel, E. W., Devenport, R. K., & Rogers, C. M. (1972). Proto-cultural aspects of chimpanzees' responsiveness to novel objects. *Folia Primatologica*, *17*, 161–170.

Mesoudi, A., & Whiten, A. (2004). The hierarchical transformation of event knowledge in human cultural transmission. *Journal of Cognition and Culture*, *4*, 1–24.

Mesoudi, A., Whiten, A., & Laland, K. N. (2006). Towards a unified science of cultural evolution. *Behavioral and Brain Sciences*, *29*, 329–383.

Nagell, K., Olguin, K., & Tomasello, M. (1993). Processes of social learning in the tool use of chimpanzees and human children. *Journal of Comparative Psychology*, *107*, 174–186.

Nielsen, M. (2006). Copying actions and copying outcomes: Social learning through the second year. *Developmental Psychology*, *42*, 555–565.

Plotkin, H. (2003). We-intentionality: An essential element in understanding human culture. *Perspectives in Biology and Medicine*, *46*, 283–296.

Reader, S. M., & Laland, K. N. (2000). Diffusion of foraging innovations in the guppy. *Animal Behaviour*, *60*, 175–180.

Richerson, P., & Boyd, R. (2005). *Not by genes alone: How culture transformed human evolution*. Chicago, IL: University of Chicago Press.

Thompson, D., & Russell, J. (2004). The ghost condition: Imitation versus emulation in young children's observational learning. *Developmental Psychology*, *40*, 882–889.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Vederhus, L., & Krekling, S. (1996). Sex differences in visual spatial ability in 9-year-old children. *Intelligence*, *23*, 33–43.

Vygotsky, L. S. (1981). The genesis of higher mental functions. In J. V. Wersch (Ed. and Trans.), *The concept of activity in soviet psychology* (pp. 144–188). Armonk: M. E. Sharpe.

Want, S. C., & Harris, P. L. (2001). Learning from other people's mistakes: Causal understanding in learning to use a tool. *Child Development*, *72*, 431–443.

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

Warrick, P. D., & Naglieri, J. A. (1993). Gender differences in planning, attention, simultaneous and successive (Pass) cognitive-processes. *Journal of Educational Psychology*, *85*, 693–701.

Whiten, A., Custance, D. M., Gomez, J. C., Teixidor, P., & Bard, K. A. (1996). Imitative learning of artificial fruit processing in children (*Homo sapiens*) and chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, *110*, 3–14.

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

Whiten, A., Horner, V., & de Waal, F. B. M. (2005). Conformity to cultural norms of tool-use in chimpanzees. *Nature*, *437*, 737–740.

Wood, D., Wood, H., Ainsworth, S., & O'Malley, C. (1995). On becoming a tutor: Toward an ontogenetic model. *Cognition and Instruction*, *13*, 565–581.

Younger, M., Warrington, M., Gray, J., Ruddock, J., McLellan, R., Bearne, E., et al. (2005). *Raising boys' achievement*. Department for Education and Skills Report: RR636.

Acknowledgments

The Economic and Social Research Council funded this research. Andrew Whiten was supported by a Leverhulme Trust Major Research Fellowship and later a Royal Society Leverhulme Senior Research Fellowship. We are grateful to Andrew Burnley for construction of the extractive tool-use box, and to all the children, parents and staff at Menzieshill Nursery, Jessie Porter Nursery, Fintry Nursery, Westfield Nursery, Castlehill Primary School, Craigiebarns Nursery and Primary School

Notes

1. All variables, except 'touched doors', had Kappa scores of .97 and above. Discrepancies in the 'touched doors' coding were discussed by coders and agreement was achieved.
2. There was no difference for the achievement score at the individual-attempt level.