Evolutionary Human Sciences (2022), **4**, e5, page 1 of 23
doi:10.1017/ehs.2022.4

RESEARCH ARTICLE

Young children spontaneously invent three different types of associative tool use behaviour

E. Reindl[1,2,3]*, C. Tennie[4], I. A. Apperly[1], Z. Lugosi[5] and S. R. Beck[1]

[1]School of Psychology, University of Birmingham, Brimingham, UK, [2]Department of Anthropology, Durham University, Durham, UK, [3]School of Psychology and Neuroscience, University of St Andrews, St Andrews, UK, [4]Department for Early Prehistory and Quaternary Ecology, University of Tübingen, Tübingen, Germany and [5]Division of Psychology, University of Stirling, Stirling, UK
*Corresponding author E-mail: eva.reindl@live.de

Abstract
Associative Tool Use (ATU) describes the use of two or more tools in combination, with the literature further differentiating between Tool set use, Tool composite use, Sequential tool use and Secondary tool use. Research investigating the cognitive processes underlying ATU has shown that some primate and bird species spontaneously invent Tool set and Sequential tool use. Yet studies with humans are sparse. Whether children are also able to spontaneously invent ATU behaviours and at what age this ability emerges is poorly understood. We addressed this gap in the literature with two experiments involving preschoolers (E1, $N = 66$, 3 years 6 months to 4 years 9 months; E2, $N = 119$, 3 years 0 months to 6 years 10 months) who were administered novel tasks measuring Tool set, Metatool and Sequential tool use. Participants needed to solve the tasks individually, without the opportunity for social learning (except for enhancement effects). Children from 3 years of age spontaneously invented all of the types of investigated ATU behaviours. Success rates were low, suggesting that individual invention of ATU in novel tasks is still challenging for preschoolers. We discuss how future studies can use and expand our tasks to deepen our understanding of tool use and problem-solving in humans and non-human animals.

Keywords: Associative tool use; Metatool use; Sequential tool use; Tool set use; Multifunctional tool use; Tool use; Problem-solving

Social media summary: Children solve tool-use tasks involving the use of two tools in combination on their own, without social learning.

Tools and technology play a dominant role in the lives of humans of all cultures. They have contributed substantially to the success of our species: our capacities for using, making and innovating tools have opened up new ecological niches (and are still doing so) and contributed to the coevolution of cumulative culture, social learning and teaching (Henrich, 2015; van Schaik et al., 1999). Today, the use and making of many tools has become so complex and/or opaque that they rely on copying of sufficient fidelity, often over extended periods (Gurven et al., 2006; Kaplan & Robson, 2002). Owing to this special role of social learning for the acquisition of many forms of tool use, researchers have investigated when, how and from whom humans, and especially children and adolescents, learn (Bjorklund & Gardiner, 2012; Esseily et al., 2016; Greif & Needham, 2011; Lancy, 2016, 2017; Lew-Levy et al., 2017, 2020; Nagell et al., 1993; Nielsen et al., 2012; Somogyi et al., 2015). These socially learned behaviours add to a phylogenetic baseline of tool use, i.e. a range of behaviours which members of the species in question can acquire *without* the need for copying – if they are at the right environmental, developmental and motivational stage. The current paper explores the

© The Author(s), 2022. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

contents of this baseline, also called the *Zone of Latent Solutions* (ZLS; Reindl et al., 2018; Tennie et al., 2009, 2020), in humans.

Some of humans' closest living relatives – chimpanzees and orangutans – also possess varied tool cultures (van Schaik et al., 2003, 2009; Whiten et al., 1999, 2001). However, in contrast to much of human tool culture, there is little evidence that the behaviours of non-human great apes (henceforth apes) require copying know-how (Bandini & Tennie, 2020; Buskell & Tennie, accepted; Motes-Rodrigo & Tennie, 2021; Reindl et al., 2018). Instead, they appear to be *socially mediated reinnovations* (Bandini & Tennie, 2017): while non-copying types of social learning (e.g. local or stimulus enhancement) can guide the individual's attention towards relevant materials and/or locations, each individual re-innovates the form of the behaviour – the know-how – on their own from their ZLS. Owing to their limited engagement in form copying, apes have been suggested to be restricted to ZLS-only cultures. If true, all behavioural forms observed in wild apes are latent solutions (Reindl et al., 2018), and strong evidence for this was provided recently (Motes-Rodrigo & Tennie, 2021). Modern humans – through their capacity for copying know-how – have managed to go beyond their ZLS, i.e. are able to learn and produce culture-dependent forms (i.e. forms which are too complex or arbitrary to be re-innovated from scratch).

By definition, cumulative culture does not arise from zero. The seemingly unbounded landscape of tools and technology modern humans have explored because of their cumulative cultural abilities rests on a second landscape which can be explored by individual learning alone: the ZLS. This second 'base' landscape is less well understood (Neldner et al., 2020; Reindl et al., 2016), but can give us important insights – via cognitive cladistics – into the tool-using abilities of our ancestors before the emergence of cumulative culture (Tennie et al., 2016). One way to investigate this landscape is through *latent solution (or baseline) tests*. These are experimental studies in which individuals naive to a cultural behaviour in question are provided with the raw materials necessary to produce the behaviour to study whether they spontaneously re-innovate the behaviour (Bandini & Tennie, 2017; Tennie et al., 2009). While Luria and Vygotsky (1930) thought that all of children's tool use stems from copying and teaching, with children's spontaneous tool use being 'practically zero' (p. 114), we now know that infants from 18 months of age can spontaneously use a tool to obtain an out-of-reach object (Fagard et al., 2014; Rat-Fischer et al., 2012) and that from at least 3 years of age children have the capacity to invent the correct solution to simple, but novel, stick tool-use problems on their own, without any immediate help from others (Neldner et al., 2020; Reindl et al., 2016). This might not be surprising, given that the ingredients for successful tool use – increasing motor skills, knowledge about object affordances, refined perception–action routines, causal cognition – are built up gradually from early infancy through social and asocial object exploration and play (Bjorklund & Gardiner, 2012; Chen et al., 2000; Greif & Needham, 2011; Kahrs et al., 2013; Lockman, 2000; Somogyi et al., 2015).

The current study aimed to add to the exploration of the contents of the human ZLS (here by studying a *WEIRD* (Westernized, Educated, Industrialised, Rich, Democratic) population (Henrich et al., 2010), but cross-cultural work is to follow) by asking whether a more complex type of tool use also lies within the human ZLS: namely *Associative Tool Use* (ATU) – the use of two or more tools in combination to achieve a goal (Shumaker et al., 2011). ATU is arguably cognitively more demanding than 'simple' tool use (i.e. the use of a single tool; Boesch, 2013): by definition, ATU involves a larger number of objects, and often consists of an increased temporal and/or spatial problem–solution distance, resulting in an increase in the processing complexity of the task and subsequently its cognitive load on working memory and other executive functions such as planning, inhibition, behavioural sequencing and decision making (Haidle, 2010; Halford et al., 1998; Hunt et al., 2013; Read, 2008). Note, however, that the adjective 'associative' in ATU should only be understood in its literal meaning ('relating to, connecting', from Latin *associatus* (= joined with)), used to describe a type of tool use in which two or more tools are used in combination. It should not be understood to imply a relationship to associative learning (nor any other learning mechanism). In the terminology by Shumaker et al. (2011), simple tool use refers to the use of a single tool, while ATU refers to the use of two or more tools in combination. The question of underlying learning mechanisms is a separate one and not alluded to by using these terms. While simple tool use is assumed to be

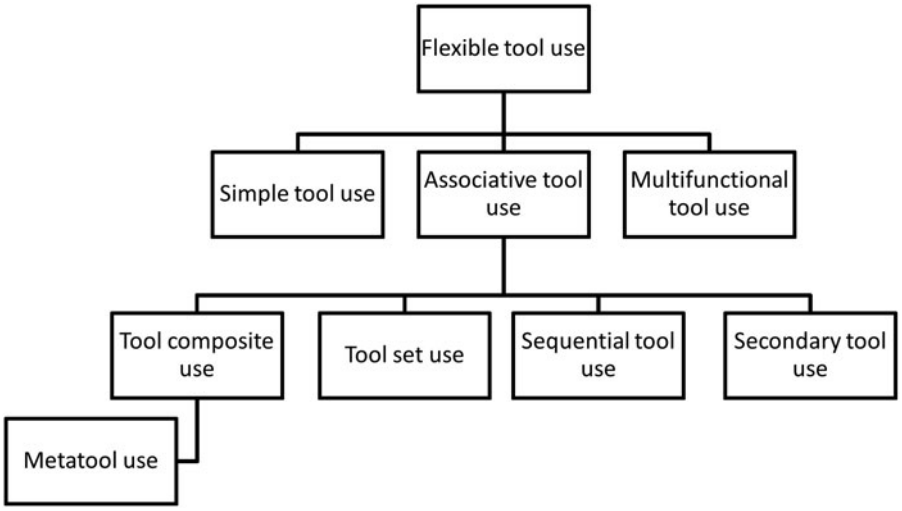

Figure 1. Classification of flexible tool use types as used in this study, based on the definitions in Shumaker, Walkup, and B. B. Beck (2011).

cognitively less demanding than ATU, this does not imply that simple tool use behaviours can always be acquired by associative learning only.

ATU can be split into several categories, according to the different ways in which tools can be used in combination. To date, there is no uniform nomenclature to describe these different forms (compare e.g. Colbourne et al., 2021; Shumaker et al., 2011; Taylor et al., 2007; Wimpenny et al., 2009). We do not aim to add to the potentially confusing terminology by introducing yet another set of definitions, and therefore refer to the nomenclature used by Shumaker et al. (2011) (Figure 1, Table 1), who have become a widely cited source in the field. We do acknowledge the potentially confusing use of the word 'associative', which many comparative psychologists are more familiar with in the context of associative learning, but hope that the above has made clear that the two concepts are not related.

We focused on three out of the four ATU types by Shumaker et al. (2011; for definitions and examples see Table 1): *Tool set use*, *Sequential tool use*, and a special case of using tool composites, *Metatool use*. We did not include Secondary tool use as previous research had shown that tool making and innovation *per se* are difficult for children (Beck et al., 2011; Nielsen et al., 2014; but see Voigt et al., 2019).

ATU has been reported in many species in which flexible (i.e. non-stereotyped) tool use has been observed (for a review, see Table S1 in the OSF repository, which is based on Shumaker et al., 2011, but also comprises research published since then). For example, Tool sets are used by wild chimpanzees and capuchin monkeys: chimpanzees use sets of two or more sticks to get access to beehives, termite mounds or ant nests – stout branches are used as pounding tools and levers to make and widen holes, and finer twigs are used on these holes as dipping sticks (Bernstein-Kurtycz et al., 2020; Shumaker et al., 2011). Capuchin monkeys use stones to pound on beehives or next to cavities, and then use sticks in these holes to probe for honey or small animals (Mannu & Ottoni, 2009). Evidence from latent solution tests with captive chimpanzees (Bernstein-Kurtycz et al., 2020), capuchin monkeys (Westergaard et al., 1997) and Goffin's cockatoos (O'Hara et al., 2021) shows that Tool set use falls into the ZLS of these species.

Metatool use is shown by wild chimpanzees in the context of nut-cracking: when stones are used as hammers and anvils, some individuals at Bossou (Republic of Guinea) use smaller stones as wedges to stabilise the anvils (Carvalho et al., 2008; Matsuzawa, 1991). In addition, wild chimpanzees and captive orangutans have been observed to use a stick to push a leaf/paper into a tree-hole or a puddle of liquid to retrieve water/juice more efficiently (Matsuzawa, 1991, cited in Sugiyama, 1997; Lehner et al., 2011;

Table 1. Associative tool use (ATU) types and their definitions according to Shumaker, Walkup, and B. B. Beck (2011) and list of animals for whom evidence for spontaneous occurrence of ATU exists

ATU type	Definition	Examples (human and non-human)	Non-human animals for which ATU type has been observed to be produced spontaneously (in the wild or in captivity)[a]
Tool set use	Two or more tools used sequentially, usually each in a different mode, to achieve a single outcome	Using a *can opener* to open a tin can followed by scooping the contents with a *spoon*. Using a *stout stick* to open up a termite mound followed by using a *thin twig* to fish for termites	Capuchin monkeys, chimpanzees, Goffin's cockatoos
Tool composite use	Two or more tools used simultaneously, usually each in a different mode, to achieve a single outcome, where the first tool is not used to manufacture the second	Using a *fork* and knife together. Standing on a *box* and using a *stick* to reach a reward	Bonobos, capuchin monkeys, chimpanzees, gorillas, long-tailed macaques, orangutans, satin bowerbirds, sea otters
Metatool use	A tool used simultaneously with a second tool to increase the efficiency of the second tool, where the first tool (metatool) acts directly on the second. The second tool could function as a tool on its own; the metatool makes it a better tool. Every metatool is a tool composite	Using *snow chains* on *car tyres* in winter. Using a *wedge* to stabilize an *anvil*	Capuchin monkeys, chimpanzees, orangutans
Sequential tool use	A tool used to acquire another tool	Using a *ruler* to reach a *key* that has fallen behind the radiator to open a box. Using a *stick* to retrieve a *longer stick* from a box to retrieve a reward from another box	Bonobos, capuchin monkeys, chimpanzees, gorillas, New Caledonian crows, Olive baboons, orangutans, rhesus macaques, rooks
Secondary tool use	A tool used to manufacture (structurally modify) another tool.	Using a *cutting machine* to produce matchsticks. Using a *stone hammer* to produce a *flake for cutting*	Capuchin monkeys

[a]Based on references in Shumaker et al. (2011) and a literature search for more recent papers, see Table S1 in the OSF repository.

Lethmate, 1982). Captive capuchin monkeys pounded chisel stones with hammer stones to break open lids, thus being more effective than when using the chisel stones alone (Westergaard et al., 1996; Westergaard & Suomi, 1994a, b). Again, the studies on captive, target behaviour naive individuals suggest that Metatool use is within the ZLS of chimpanzees and capuchin monkeys, i.e. does not need to be copied by others to be acquired.

There are no reports of Sequential tool use in wild non-human animals. Instead, Sequential tool use problems have been created as novel laboratory tasks to investigate problem-solving and causal cognition in various species and date back to Köhler's (1921) work with chimpanzees. Research has shown that New Caledonian crows, all four species of great ape, capuchin monkeys and baboons are able to spontaneously and unaidedly solve Sequential tool use tasks, in which short sticks need to be used to retrieve longer sticks, which in turn can be used to access a reward (Anderson & Henneman, 1994; Bolwig, 1963; Martin-Ordas et al., 2012; Mulcahy et al., 2005; Taylor et al., 2007, 2010; Wimpenny et al., 2009).

In contrast to the non-human animal literature, the low number of studies on ATU in humans is striking. Köhler (1921) remarked that researchers were still facing terra incognita regarding understanding children's flexible tool behaviors, and 90 years later, researchers still pointed out the only patchy understanding of the development of tool use in children (Greif & Needham, 2011). While in the meantime a number of studies on the origins and development of simple tool use, tool making and tool innovation in children have been conducted (Barrett et al., 2007; E. Bates et al., 1980; Bechtel et al., 2013; S. R. Beck et al., 2016; Breyel & Pauen, 2021; Brown, 1990; Chappell et al., 2013; Chen et al., 2000; Deák, 2014; Gönül et al., 2018; Keen, 2010; Lew-Levy et al., 2021; McCarty et al., 1999, 2001; Neldner et al., 2020; Pauen & Bechtel-Kuehne, 2016; Piaget, 1952; Rat-Fischer et al., 2012; Rawlings et al., 2021; Reindl et al., 2016; Voigt et al., 2019; Willatts, 1984), there are only few studies on *associative* tool use in children, and all of them focus on only one type: Sequential tool use (Alpert, 1928; Matheson, 1931; Metevier, 2006). Köhler (1921) remarked that most tool-use behaviours in adults have become 'mechanised', i.e. are carried out with ease and so questions about their ontogenetic and phylogenetic origins might not present themselves as very salient. Yet research on ATU allows us to (a) further illuminate the extent of the human ZLS, (b) improve our understanding of the evolution and development of tool-using and innovative skills and their underlying cognitive processes (using cognitive cladistics) and (c) provide insightful comparisons to the existing studies involving ATU in non-humans.

Alpert (1928) and Matheson (1931) adapted Köhler's (1921) tool-use tasks for use with preschool children. In their studies, children were separated from a reward by a railing set up in the testing room. To obtain the reward, children had to use a stick lying on their side of the railing to rake in a longer out-of-reach stick beyond the railing, which they could then use to fetch the reward. Children below 2.5 years of age did not solve this task; only 10 out of the 28 tested 2.5- to 4.5-year-olds (36%) spontaneously solved it (Matheson, 1931). In a more recent study, 3-year-olds completed two tasks involving the use of a single tool before attempting an ATU task (Metevier, 2006): In the 'tube task', participants had to use a stick to push a toy out of a tube. In the 'rake task', they had to use a rake to obtain an out-of-reach toy on the table. Children found these tasks rather easy, indicated by success rates above 75%. In the 'combination task' (ATU), children first had to use the rake to obtain an out-of-reach stick on the table, and subsequently use this stick to push a toy out of the tube. Success rates were low (25–37%), indicating that 3-year-olds struggled with Sequential tool use even though they readily solved the individual components beforehand. These studies suggest that Sequential tool use is challenging for children up to at least the age of 4 years, both when having to solve such tasks spontaneously and given previous experience with individual elements of the tasks. However, no research seems to have been carried out yet on Metatool and Tool set use abilities in children.

The current study aimed to investigate whether young children would be able to solve three types of ATU tasks (Tool set use, Sequential tool use and Metatool use) on their own in a latent solution test, i.e. without immediate social learning (Experiment 1). In Experiment 2 we further compared the

difficulty of Tool set and Sequential Tool use using apparatuses that were more comparable between these ATU types. The age range of our samples was slightly higher than in a previous latent solution test on simple tool use (Reindl et al., 2016) as it was assumed that ATU would pose greater demands on executive functions, such as working memory, which are still developing during childhood (Garon et al., 2008). However, note that increasing the age range also increases the amount of previous cultural knowledge that children bring to the experiment and thus the probability that successful children can use previously acquired cultural knowledge to solve the tasks.

Experiment 1

We investigated whether 3.5- to 4-year-old children would be able to spontaneously solve three types of ATU tasks: Tool set use, Metatool use and Sequential tool use. We created six tasks (two tasks per ATU type); four were based on ATU behaviours observed in wild or captive animals (one Tool set use task, both Metatool use tasks, one Sequential tool use task; see below) and two were new creations. Each child was administered three tasks in a single session, with one task from each ATU category (task order counterbalanced). During data collection, we noticed that one of our Metatool use tasks (*Anvil prop*) had a design flaw as most children (eventually 71%) were able to solve the task in a way not intended by the task design, i.e., without ATU. This task is described in the Supplementary Material but is not included in the analysis.

Methods

Participants

We tested 66 children (31 boys) between 3 years 6 months and 4 years 9 months (dates of birth were known for 64 children: mean ± SD, 4 years 1 month ± 3.88 months) in seven nurseries and a Science museum in Birmingham, UK, between March and July 2014. No *a priori* sample size calculation was carried out for either Experiment 1 or 2; instead, the goal was to test as many children as possible during the time window available for this project. The ethnic composition of the sample was 65.2% White, 21.2% Black and 13.6% Asian. Participants were recruited through letters sent to parents (for children tested in nurseries) and via advertisements on the museum website and social media (for children tested at the museum). Ethical approval for both experiments was granted by the University of Birmingham, UK, STEM Ethical Review Committee.

Materials

Drawings of the tasks are displayed in Figure 2 (for photos of the tasks, see Figure S1). For space reasons, tasks are only briefly described here; a full description including animations on how each task could be solved can be found in the Supplementary Material. For each task, the apparatus(es) were positioned in front of the child and all freely accessible tools were placed between the task and the child. Rewards were either stickers placed in containers or other target objects which could be exchanged for a sticker after completion of the task.

Tool set use – Open and probe. This task was based on the use of Tool sets by chimpanzees to open beehives, ant nests or termite mounds (see above). The goal was to insert a short, sturdy stick into a tube and pierce a tin foil barrier, and then to use a long pipecleaner with hook and loop fastener wrapped around both ends to reach through the hole in the barrier and retrieve a target object from the back end of the tube.

Tool set use – Push and hook. This task was novel. The goal was to insert a short, sturdy stick into the side or top openings of a T-shaped apparatus to move a barrier inside the apparatus, and then to insert a rope with a hook at its end through the top of the apparatus to fish for a bucket containing a reward (this second step was similar to the hook task in Chappell et al., 2013).

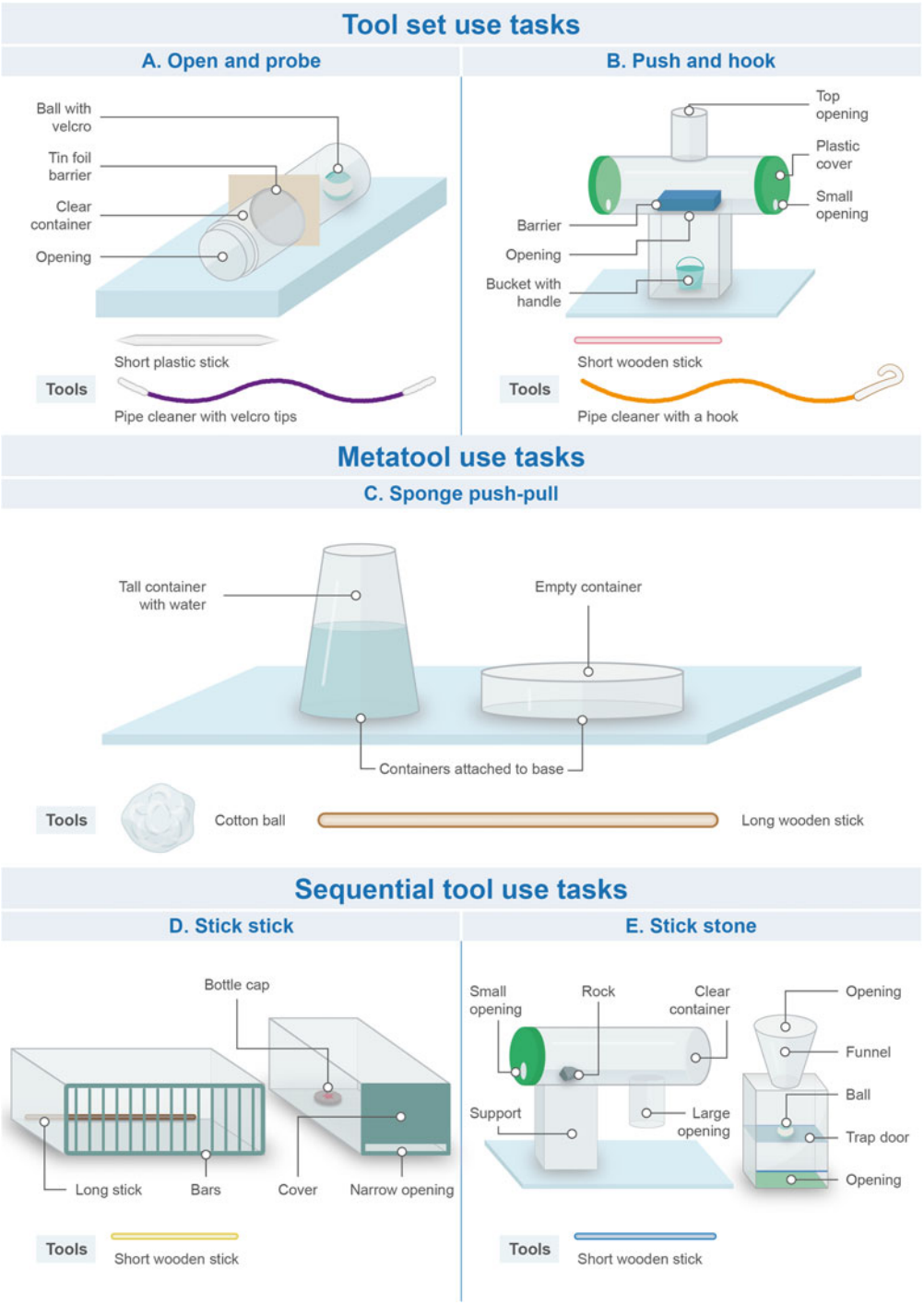

Figure 2. Materials used in Experiment 1.
Note: Anil prop task not shown as it was excluded from the analysis due to design failure. Figure drawn by Nuria Melisa Morales García.

Metatool use – Sponge push–pull. This task was based on observations of chimpanzees and orangutans using a stick (metatool) to push a leaf/paper towel into a tree-hole or puddle of liquid to retrieve water/juice more efficiently (Matsuzawa, 1991, cited in Sugiyama, 1997; Lehner et al., 2011; Lethmate, 1982). Children were presented with a transparent tube filled with 500 ml of water and a smaller container placed next to it. The goal was to fill the small container with water. Available tools were a wooden stick and a ball of cotton wool. While the stick could be used on its own to solve the task – by dipping it into the tube repeatedly to extract water, this method was inefficient. Instead, the stick use could be improved by the wool as a metatool: children could first drop the wool into the tube so that it could absorb water, after which they could use the stick to retrieve the wool. Note that because of this retrieval action (stick used to retrieve wool from the bottle), the Sponge push–pull task might resemble a Sequential tool use task, which by definition involves the use of a tool to retrieve another tool. However, the retrieval action is just a feature of this particular task. In Metatool use, a tool (here the stick) is used to carry out a target action (here to retrieve the water from the bottle) and the metatool (here the wool) enhances the efficiency of the target action. In contrast, the defining feature of Sequential tool use is that a tool is used to retrieve a second tool in a first step, and then this second tool can be used in a second step to carry out the target action. More information about this differentiation can be found in the Supplementary Material.

Sequential tool use – Stick stick. This task was based on an apparatus used for studying Sequential tool use in New Caledonian crows (see figure 1 in Taylor et al., 2007). The goal was to insert a short stick into a box to rake in a longer stick, which could then be used on another box to obtain an-out-of-reach sticker.

Sequential tool use – Stick stone. At the time we designed this experiment, this task was novel. Note, however, that a similar design has subsequently been used with New Caledonian crows (Gruber et al., 2019). Metevier (2006) remarked that in most Sequential tool use studies the tools were of the same type (usually sticks of different sizes) and used in a similar fashion (e.g. raking). Therefore, we created a task in which two different kinds of tools (a stick and a stone) had to be used in different ways (pushing, dropping). The goal was to insert a short stick into an apparatus to push out a small stone. The stone could then be dropped into the top of a second apparatus where it would activate a trapdoor and release a target object.

Design and procedure

For reasons of practicability, we tested children only once and thus only administered a subset of the tasks (a warm-up game and three tasks) to each child. The combination and order of the tasks were counterbalanced. Across children, every task was presented 24 times, occurring eight times in each of the three positions. Participants were tested individually by the same female experimenter (E.R.) and were sitting at a table or on the floor, perpendicular to the experimenter. A warm-up game was used to familiarise children with breaking and modifying material within the experimental session. This was important as the Anvil prop task required children to apply physical force and break the plastic nut. For the test phase, children were presented with three semi-randomly chosen tasks, one from each ATU category. Tasks were presented as a game to the children in which they could win stickers. Materials were placed in front of the participant, with the tools lying between the apparatus(es) and the child. Children were told the goal of the task, e.g. 'to get this orange ball out of the bottle' (Open and probe) and that they could 'use anything here on the table/floor' to solve the task, but they were never told that they had to use the tools to solve the tasks and they only received general encouragement. Children had 3 min to solve each game. Trials ended when children obtained the target, when time was over or if children refused to play. When one trial ended, the experimenter cleared the table/floor and fetched the next task. Children were rewarded with stickers regardless of success. Children were never shown the correct solution of a task if they were unsuccessful. Total testing time was ~15 min.

Scoring and analysis

Children's behaviour was live-coded and coding was double-checked offline by E.R. using videos from nine children (14% of the sample; note that videos were not available for all participants). For each task, we scored whether children picked up the tool(s), used the tool(s) in the manner intended by the experimenter, engaged in ATU (i.e. whether children used both tools to solve the task in a manner that was intended by us) and whether they solved the task following ATU (*Correct success*; i.e. whether children succeeded after having used both tools in the way intended by us). We scored *Incorrect success* whenever children solved the task in a way that was not intended by us. To obtain inter-observer reliability, 31% of the valid trials (i.e. 50 trials) were live-coded by a second rater who was present during the experiment and asked to code our two main variables, ATU and Correct success. Inter-rater agreement for both variables was perfect (Cohen's $k = 1.000$).

Each of the 66 children participated in three tasks, resulting in 198 trials. From these, we excluded all of the Anvil prop trials ($n = 33$) and a further three trials owing to experimenter error (Tube task, $n = 2$) or because the child became upset (Tube task, $n = 1$), resulting in a final number of 162 valid trials across five tasks.

Analyses were carried out in R version 4.0.3 (R Core Team, 2020). To address our main question – whether children were able to spontaneously engage in ATU to solve novel problems – we carried out descriptive analyses for each task, investigating success rates. In exploratory analyses, we examined whether the two tasks within the Tool set and Sequential tool use types were of comparable difficulty by conducting chi-square analyses of children's rates of ATU and Correct success. We also examined whether ATU type affected children's ATU and Correct success using two Generalised Linear Mixed Models (GLMMs; Baayen, 2008) with binomial error structure and logit link function (McCullagh & Nelder, 1989) implemented by the *glmer* function of the R package *lme4* (D. M. Bates et al., 2013). In both models, age in months (z-transformed to a mean of zero and a standard deviation of 1) was entered as control variable and ATU type as the predictor variable. Participant ID was included as a random effect to account for the fact that each child contributed several datapoints. Model stability was assessed by comparing the estimates obtained from the model based on all data with those obtained from models with the levels of the random effects excluded one at a time (Nieuwenhuis et al., 2012). There were no issues with model stability (Tables S7 and S8). As an overall test of the effect of ATU type we compared each full model with a null model lacking the test predictor but keeping age and the same random effects structure as the full model (Forstmeier & Schielzeth, 2011) using a likelihood ratio test (Dobson, 2002). The data and script for Experiments 1 and 2 can be found on the OSF website: https://osf.io/d3pz5/?view_only=ba368e675f324a56a7cad600d6c39581.

Results

Table 2 provides an overview of the ATU and Success scores of the tasks. For a detailed breakdown of how children attempted to solve the tasks and at which points in the process children got 'stuck', see the Supplementary Material. For each task, two or more children were able to spontaneously, i.e. unaidedly and within a very short time span of 3 min, engage in ATU and solve the task. The double-case ZLS standard requires that for relatively less complex behaviours, which have some (low) probability of occurring by chance, two independent, naive individuals must demonstrate the behaviour. According to this standard, this allows ZLS researchers to be confident enough to conclude that the behaviour was re-innovated from the species' ZLS, rather than being a chance occurrence (Bandini & Tennie, 2017; see also Supplementary Material for more details). This is the case here. In addition, the materials and problems in our tasks were rather unfamiliar to the participants. Together, these results support the view that simple ATU behaviours as presented here are within the human ZLS and do not rely on social learning to be acquired.

We examined whether the two tasks within the Tool set use type were of comparable difficulty. We found that the Open and Probe task had higher rates of ATU than the Tube task ($\chi^2(1) = 9.98$, $p = 0.001$) as well as higher Correct success rates ($\chi^2(1) = 12.10$, $p < 0.001$), indicating that the Tool

Table 2. Number (and percentage) of valid trials in which ATU, correct success and incorrect success were scored in Experiments 1 and 2

	Experiment 1				Experiment 2			
ATU type	Task ($n_{\text{valid trials}}$)	Associative tool use	Correct success	Incorrect success	Task ($n_{\text{valid trials}}$)	Associative tool use	Correct success	Incorrect success
Tool set use	Open and probe (33)	23 (69.8%)	18 (54.5%)	2 (6.1%)	Box (40)	6 (15.0%)	4 (10.0%)	4 (10.0%)
	Tube task (30)	8 (26.7%)	3 (10.0%)	1 (3.3%)	Tube (38)	9 (23.7%)	7 (18.4%)	0 (0%)
Metatool use[a]	Sponge push–pull (33)	8 (24.2%)	7 (21.2%)	0 (0%)				
Sequential tool use	Stick stick (32)	2 (6.3%)	2 (6.3%)	6 (18.7%)	Box (38)	11 (28.9%)	11 (28.9%)	1 (2.6%)
	Stick stone (34)	2 (5.9%)	2 (5.9%)	4 (11.8%)	Tube (39)	2 (5.1%)	2 (5.1%)	0 (0%)
Multifunctional tool use					Box (40)	11 (27.5%)	11 (27.5%)	0 (0%)
					Tube (40)	31 (77.5%)	30 (75.0%)	0 (0%)
N_{total}	162 trials from 66 children	44 (27.2%)	32 (19.7%)	13 (8.0%)	235 trials from 119 children	70 (29.8%)	65 (27.6%)	5 (2.1%)

[a]Data for Anvil prop not shown here as task was removed from the analysis owing to design failure.

set use tasks were not of equal difficulty. We also compared the two tasks within the Sequential tool use type and found them to be of equal difficulty regarding the ATU and Correct success rates (Fisher's exact tests, both $p = 1.000$). As we noticed during data collection that there was a design flaw with the Anvil prop task (the majority of children solved the task in a way not intended by us and not involving ATU), we excluded Anvil prop from all analyses. Therefore, Metatool use was not included in this analysis.

We examined whether ATU type affected children's ATU and Correct success rates. The models comprised 144 trials from 63 children (cases of Incorrect success were removed from the analysis). ATU type had a significant positive effect on ATU rates ($\chi^2(2) = 30.70$, $p < 0.001$). Specifically, Tool set use yielded significantly higher ATU rates than Sequential tool use ($p < 0.001$) and Metatool use ($p = 0.030$; note that Metatool use only consisted of the Sponge push–pull task). The odds for children in the Tool set use condition engaging in ATU over the odds of children in the Sequential tool use condition were 15.01 (95% CI [6.16; 100.38]). The odds for children in the Tool set use condition scoring ATU over the odds of children in the Metatool use condition were 4.05 (95% CI [1.57; 14.79]). Performance in the Sequential tool use and Metatool use tasks was equally low ($p = .135$; Table S7). ATU type also had a significant positive effect on Correct success rates ($\chi^2(2) = 15.08$, $p < 0.001$). Specifically, Tool set use yielded significantly higher Correct success rates than Sequential tool use ($p = 0.003$). The odds for children in the Tool set use condition scoring Correct success over the odds of children in the Sequential tool use condition were 8.98 (95% CI [2.86; 35667.67] – but note the large uncertainty). No other comparisons were statistically significant (Table S8).

Discussion

Experiment 1 investigated whether 3.5- to 4-year-old children were able to independently re-innovate how to use two tools in different combinations to solve several problem-solving tasks. To our knowledge, this is the first time several ATU types have been investigated in children in a single study. We found that children succeeded in all three tested types of ATU individually, without the need for social learning immediate to the experimental context (note that past social learning could still matter). However, success rates were low, with only one task (Open and probe (Tool set use)) having a success rate of more than 50%, suggesting that the individual invention of these behaviours is – while not impossible – quite challenging for 3- to 4-year-old children. Sequential tool use was especially challenging (~6% Correct success rate). This might be due the fact that the Sequential tool use tasks involved two apparatuses, which might have increased general task difficulty compared with the other ATU types. However, despite this, our tasks were still relatively easy versions of Sequential tool use: we presented only one initial tool (rather than a choice of tools), with which only one other tool could be retrieved. The cost of retrieving the second tool was relatively small and the two apparatuses were in proximity (in contrast, in some studies involving New Caledonian crows, apparatuses are positioned opposite each other (Wimpenny et al., 2009) or completely out of the sight of the others (Gruber et al., 2019) so that subjects have to keep the necessary information in their short-term memory, providing strong evidence of New Caledonian crows mentally representing stages of the problem). The different types of ATU were not equally easy, with Tool set use being the easiest type of ATU, and Sequential tool use possibly the hardest. This could be interpreted as a first hint at a potential 'cognition-based hierarchical organization' (Shumaker et al., 2011, p. 21) of ATU. Yet, our results should be treated with caution until future research has been carried out.

Experiment 2 followed up on the question whether Sequential tool use actually posed more cognitive demands than Tool set use tasks or whether the finding from Experiment 1 was caused by task-specific effects. In Experiment 2, we used the *same* apparatus to investigate children's spontaneous engagement in Tool set and Sequential tool use behaviours, removing the confound of task-specific influences on performance. To also investigate potential age effects on performance, we tested a slightly wider age range (3–6 years). Experiment 2 also involved a condition requiring Multifunctional tool use (i.e. using a single tool in different functions) to contrast against Tool set and Sequential tool use.

Experiment 2

We used two apparatuses (*box apparatus* and *tube apparatus*; Figure 3), which could be presented in two ATU types (Tool set use, Sequential tool use). Presenting the ATU versions on two apparatuses and comparing children's performance across them was a first step towards disentangling effects based on the cognitive demands of the respective ATU type from more task-specific demands (e.g. differences in transparency of the apparatuses, or how easily alternative solutions could be found). A third presentation mode was Multifunctional tool use (described below, but not a major focus of this paper). Children were randomly assigned to one of these three tool-use types (Tool set use, Sequential tool use, Multifunctional tool use). Participants completed two trials within their tool-use type, one with the box and one with the tube apparatus (order counterbalanced).

Methods

Participants

The final sample included 119 children (53 boys) between 3 years 0 months and 6 years 10 months (Mean ± sd: 5 years 0 month ± 12.38 months) tested in nurseries, schools and a Science Museum in Birmingham, UK ($n = 95$) and a nursery in Jena, Germany ($n = 24$) between January and March 2015. The ethnic background of the sample was mainly White ($n = 109$; 91%), nine children (7%) were Black, and one (2%) was Asian. Three additional children were tested but removed from the sample because they were below the age of 3 ($n = 2$) or because of interference from nursery staff ($n = 1$).

Material

We used two apparatuses which could be administered in each of the three tool-use versions (Tool set, Sequential, Multifunctional tool use; Figures 3 and S7). For space reasons, tasks are only briefly described here; a full description including animations can be found in the Supplementary Material.

Box apparatus – Tool set use. This task was similar to the Open and Probe task of Experiment 1. The goal was to insert a short, sturdy stick into the box to make a hole in a tin foil barrier and to then use a pipecleaner to reach through the hole to retrieve a reward.

Box apparatus – Sequential tool use. In this version, there was no barrier inside the box. The tools were the same as in Tool set use (short stick, pipecleaner). The pipecleaner was placed inside the box at the place where the barrier was positioned in Tool set use. The task required using the stick to retrieve the pipecleaner from the box, followed by using the pipecleaner to retrieve the container with the reward.

Box apparatus – Multifunctional tool use. As in Tool set use, the interior of the box was divided by a tin foil barrier. However, only one tool was available, combining the functions of the tools in the Tool set use version: a wooden stick covered with hook and loop fastener at both ends, which could both pierce the tin foil barrier and reach and retrieve the container.

Tube apparatus – Tool set use. The task was similar to the Push and hook task of Experiment 1. The goal was to insert a stick into the apparatus to remove a cloth barrier by pushing or pulling it via the top or side openings, followed by inserting a rope into the top to hook a bucket with a sticker.

Tube apparatus – Sequential tool use. The tools were the same as in Tool set use, but there was no cloth barrier inside the apparatus. The rope was placed inside the box, close to the side facing the participant. The task required using the stick to retrieve the rope via the side openings, after which the rope could be inserted into the top to hook the bucket.

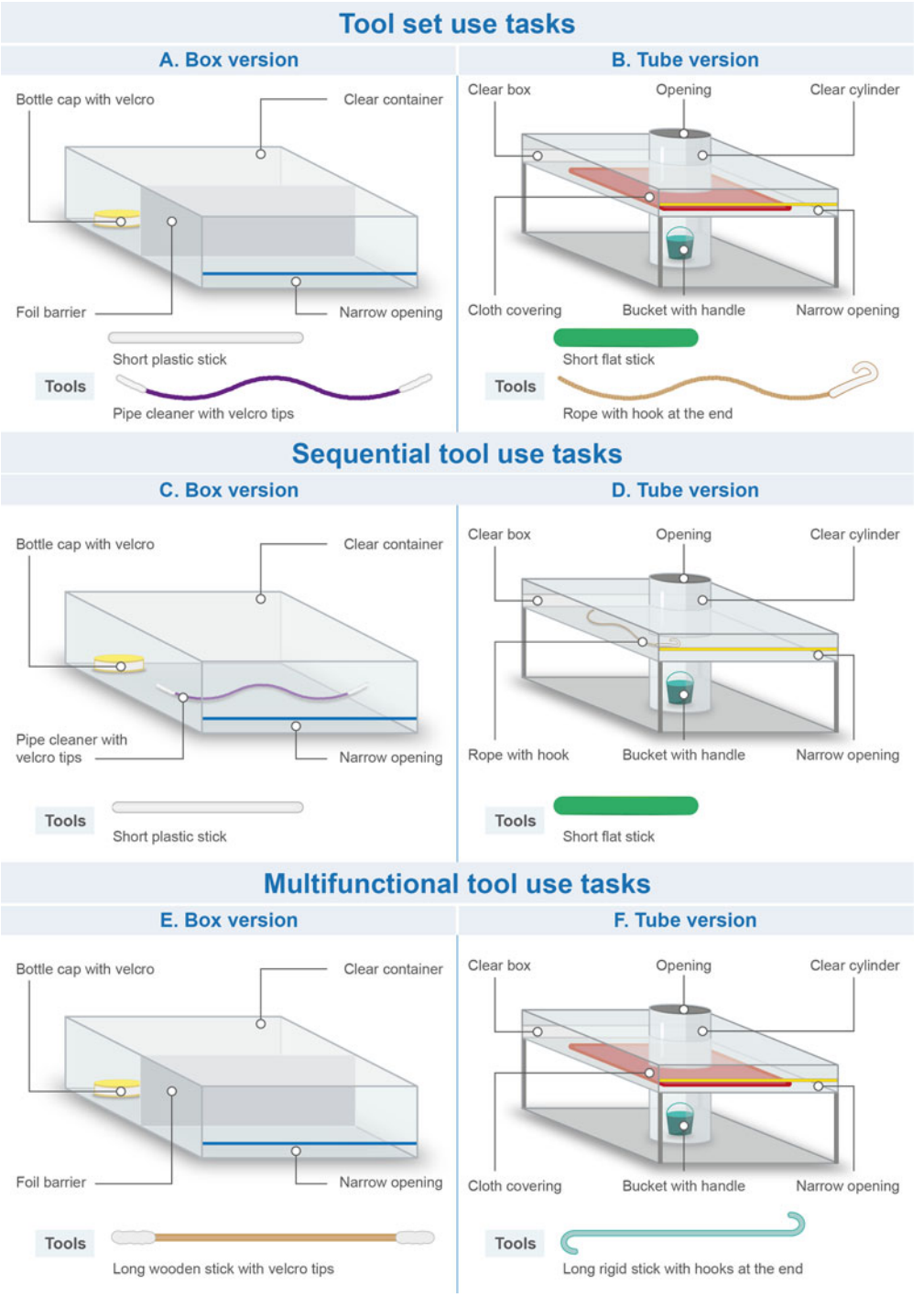

Figure 3. Materials used in Experiment 2. Figure drawn by Nuria Melisa Morales García.

Tube apparatus – Multifunctional tool. The setup was as in Tool set use, but there was only one tool available combining the functions of the two tools in Tool set use: a plastic stick with two hooks facing in opposite directions which could both remove the barrier and hook the bucket.

Design and procedure

The procedure was similar to Experiment 1, with the difference that here participants were given 4 min to solve the task (for a full description of the procedure, see Supplementary Material).

Scoring and analysis

Children's behavior was live-coded. We recorded whether the tool(s) associated with each task were picked up and used in the manner intended by the experimenter. For Sequential tool use and Tool set use we again scored ATU and Correct success. In the Multifunctional tool use task, we scored whether the tool was used in both its intended functions (equivalent to ATU) and Correct success. We also scored instances of Incorrect success for all tasks. To obtain inter-observer reliability, 84% of the valid trials were coded by a second rater (68% of these trials were coded live, the rest from video). Interrater agreement for Correct success was very good (Cohen's $k = 0.988$).

All 119 children participated in two tasks, resulting in 238 trials. Three trials had to be removed from the analysis: in the Tool set use version of the Tube task, one trial had to be excluded as the cloth did not fully cover the hole and another trial had to be excluded because the child cried. In the Sequential tool use version of the Box task, one trial had to be excluded because the child tipped the box. We first carried out descriptive analyses. Then we examined whether the two versions within each tool use type were of comparable difficulty regarding ATU and Correct success by conducting McNemar tests (with continuity correction). To investigate the effects of condition and age on ATU and Correct success, we ran two GLMMs with binomial error structure and logit link function. Age in months (z-transformed) and condition were entered as predictor variables, and participant ID was included as a random effect.

Results

Table 2 provides an overview of the ATU and Success scores of the tasks (Table S9 provides the same overview split by age groups). Out of the 116 participants who completed two valid trials, 60 children (52%) were unsuccessful in both tasks, 44 children (38%) had one of the tasks correct and only 12 children (10%) were able to succeed in both tasks, suggesting that the tasks were challenging for the children. Despite the low success rates, especially in the youngest age group (Table S9), children as young as 3 years of age were able to independently re-innovate Tool set and Sequential tool use. The low success rates could not be explained by a lack of motivation as children were generally found to be interested in the tasks and interact with the tools: In each condition, at least 82% of the children picked a tool up (see Tables S10–S15). A more detailed overview of how children attempted to solve the tasks can be found in the Supplementary Material.

In the Tool set use condition, there was no significant difference between the box and tube versions for either ATU ($p = 0.579$) or Correct success ($p = 0.505$). Both versions were very challenging, partly owing to task-specific reasons that are not necessarily tied to the ATU nature of the tasks: in the box version, while most children correctly inserted the stick into the box, many did not pierce the tin foil barrier – either because they tried to reach the reward by pushing the stick over the barrier or because they failed to pierce the foil despite trying (Figure S6). Note that this difficulty in piercing the tin foil did not occur in the Open and probe task of Experiment 1. This was possibly because (a) it was not possible to bypass the tin foil barrier in the Open and probe task and (b) in Experiment 2 we used a larger piece of tin foil, which might have resulted in the foil being held in the cardboard frame more loosely, making it harder to break it. Similarly, in the tube version, many children accidentally dropped the stick into the tube or failed to remove the cloth barrier despite trying (Figure S7). If these design

aspects made the successful use of the first tool of the tool set (i.e. the stick) difficult, it means that Experiment 2 could potentially underreport children's ATU capacities in a Tool set use context.

In the Sequential tool use condition, the box version yielded significantly higher rates of ATU ($p = 0.027$) and Correct success ($p = 0.016$) than the tube version. This was possibly because many children did not notice the rope inside the tube apparatus and/or accidentally dropped the stick into the tube, which posed additional challenges to the tube version (see the Supplementary Material for a longer discussion).

In the Multifunctional tool use condition, the tube version revealed significantly higher rates of 'ATU' (i.e. using the tool correctly in both functions) and Correct success than the box version (both $p < 0.001$). In the box version, the challenges were similar to the box version of the Tool set use condition: many children either failed to break the tin foil barrier despite attempting and/or tried to reach the reward by pushing the tool over the barrier. The tube version, however, was easier in the Multifunctional tool use condition as it was not possible to drop and lose the tool in the tube (owing to its length) and because it seemed to be easier to remove the cloth barrier. In sum, the challenges the children faced were not only determined by the tool use condition, but also in large part by task-specific aspects.

The GLMMs consisted of 235 observations from 119 children. Regarding the model with ATU as a dependent variable, age and condition together explained the data significantly better than a null model only consisting of the intercept ($\chi^2(3) = 47.14$, $p < 0.001$). Condition had a significant positive effect on ATU rates ($\chi^2(2) = 33.98$, $p < 0.001$; Table S16). While Tool set and Sequential tool use did not differ in their ATU rates ($p = 0.998$), they were both significantly more difficult than Multifunctional tool use (both $p < 0.001$). The odds for children in the Tool set use condition scoring ATU over the odds of children in the Multifunctional tool use were 0.16 (95% CI [0.06; 0.32]), i.e. the odds of scoring ATU in the Tool set use condition were decreased by 84% compared with the Multifunctional tool use condition. The odds for children in the Sequential tool use condition scoring ATU over the odds of children in the Multifunctional tool use were also 0.16 (95% CI [0.05; 0.34]), i.e. also decreased by 84%. Age had a significant, positive effect on ATU rates ($\chi^2(1) = 14.68$, $p < 0.001$). With each month increase, there was an increase in the odds of scoring ATU of 1.85 (95% CI [1.37; 2.77]).

This pattern of results was similar when using Correct success as dependent variable. Age and condition together explained the data significantly better than a null model only consisting of the intercept ($\chi^2(3) = 55.51$, $p < 0.001$). Condition had a significant positive effect on Correct success rates ($\chi^2(2) = 37.52$, $p < 0.001$; Table S17). As with ATU, the Correct success rates in the Tool set and Sequential tool use conditions did not differ ($p = 0.635$) but were significantly lower than in Multifunctional tool use (both $p < 0.001$). The odds for children in the Tool set use condition scoring Correct success over the odds of children in the Multifunctional tool use were 0.11 (95% CI [0.03; 0.23]), i.e. the odds of scoring Correct success in the Tool set use condition were decreased by 89% compared with the Multifunctional tool use condition. The odds for children in the Sequential tool use condition scoring Correct success over the odds of children in the Multifunctional tool use were 0.17 (95% CI [0.06; 0.33]), i.e. they were decreased by 87%. Age had a significant, positive effect on Correct success rates ($\chi^2(1) = 21.11$, $p < 0.001$). With each month increase, there was an increase in the odds of scoring Correct success of 2.20 (95% CI [1.58; 3.51]).

Discussion

Experiment 2 investigated children's performance in Tool set and Sequential tool use tasks and compared it with their performance in Multifunctional tool use tasks. In contrast to Experiment 1, in which the tasks of different ATU types differed in both ATU type and in the apparatus and task design, in Experiment 2 the different conditions were more comparable as they were administered using the same kinds of apparatuses (a box and a tube apparatus, each of which could be presented in each of the three tool use versions).

Supporting the findings from Experiment 1, the results showed that children from 3 years of age were able to independently invent how to use two tools in different combinations to solve these

tasks, even when given no familiarisation phase, only 4 min per task and no opportunity for social learning. Despite this, the results again showed that spontaneously re-innovating ATU in novel tasks is difficult for preschoolers, at least under testing situations as used here. We found no evidence that Tool set and Sequential tool use differed in difficulty. However, both tasks were significantly more difficult than the Multifunctional tool use tasks, i.e. tasks which had the same setup as the Tool set use versions but differed in the number of tools provided (a single tool that could be used in two modes). Success rates increased with age over the tested range from 3 to 6 years. A more fine-grained analysis of solution approaches showed that performance was not only affected by the specific tool-use demands of a task (i.e. whether Sequential tool use or the use of a Tool set was required), but also to a large part by the idiosyncratic features of each task, despite the use of the same apparatus across ATU versions.

General discussion

The aim of the current study was to address the gap in the developmental literature on the emergence and spontaneous invention of ATU behaviours in human children. This study investigated children's performance in tasks involving Sequential tool use (using a tool to get a tool), Metatool use (using one tool to improve a second tool) and Tool set use (using two tools to achieve a single outcome) and showed in two experiments that children from 3 years of age can spontaneously invent solutions to all three of these ATU types individually, i.e. without the need for immediate social learning. These findings suggest that ATU behaviours lie within the human ZLS, as they have also been shown to do for all species of great apes and some other non-human primate and bird species (Table 1 and Table S1 in the OSF repository). The presence of spontaneous ATU in humans and non-human great apes suggests that the last common ancestor of humans and apes, living ~13 million years ago (Stewart & Disotell, 1998), was probably also able to engage in ATU without the need for copying know-how.

We minimised the possibility that children could draw on directly relevant cultural knowledge by using novel tasks that children were unlikely to have encountered before. This does not mean that children did not draw on more general knowledge, e.g. about affordances and the physical properties of the materials involved (e.g. ropes, hooks, pipecleaners). Assuming that the novelty of the tasks and the participants' relatively young age were effective in minimising children's ability to use specific cultural knowledge, the results suggest that all of the tested ATU types potentially lie within the human ZLS, i.e. at least some types of ATU can be invented without social learning.

Our experiments still need to be replicated in other, non-Westernised cultures to allow for a stronger conclusion about the human ZLS. In addition, future studies could attempt to test even younger children. The age range tested in the current study was slightly higher than what was used in a previous study on simple tool use (Reindl et al., 2016) as it was assumed that ATU would demand greater executive function skills (especially working memory), which are still developing in children (Garon et al., 2008; Reindl et al., 2021).

Across both experiments, success rates were low, supporting previous findings that ATU is challenging for preschoolers (Alpert, 1928; Matheson, 1931; Metevier, 2006). Experiment 2 showed that children's performance still increased with age. Our tasks were purposefully created to be challenging. Children were given only a short time to explore and attempt the tasks, which might have decreased the ecological validity of the task and artificially limited children's tool-using skills. In recent years evidence has accumulated that longer testing times will result in higher success rates in children (Breyel & Pauen, 2021; Voigt et al., 2019). Thus, it is possible that when administering our tasks with a longer time window, more children will find solutions. To increase the ecological validity of such studies, future projects should provide longer testing times and more possible solutions, and arguably also administer tasks to pairs of participants instead to individuals only (see Gönül, Hohenberger, Corballis, & Henderson, 2019; Reindl & Wronski, 2022).

How do children's low success rates compare with the success rates reported so far for non-human animals? Based on the data available to us (Table S1 in the OSF repository), we calculated the percentage of successful participants for all previous ATU experiments for which both the number of tested

individuals and the number of successful individuals were reported (note that we did not include studies listed in the 'questionable cases' sheet, e.g. those in which individuals received training or had considerable prior experience with parts of the task), resulting in 55 entries across species and ATU types with which we could compare our results (Table S1 in the OSF repository, sheet 'Success rates'). Success rates vary greatly, even within a single species and ATU type. This is not surprising as studies differed in tasks used, testing time, and in whether alternative solutions to the task were available. However, it seems that the highest success rates were produced in Sequential tool use studies – for all four great ape species, capuchin monkeys, rhesus macaques, baboons and rooks the percentage of successful individuals was substantially greater than 50% (but note that the sample sizes were extremely small) – with the exception of New Caledonian crows, whose success rate (14%) is similar to children's average success rate in the current study (12%). However, it should be noted that these numbers could misrepresent animals' abilities owing to publication bias and selective testing (i.e. in some studies only those individuals who had solved simple tool use problems were able to advance to ATU tasks, e.g. Metevier, 2006). The task which resembles most closely an already existing task is our Stick stick task (Sequential tool use, Experiment 1), as it was based on Taylor et al.'s (2007) task for New Caledonian crows (see also Taylor et al., 2011; Wimpenny et al., 2009). While the success rate in these three bird studies is at 100% and thus in stark contrast to the results reported here for children, one has to note that the crows had prior experience with one or more parts of the tasks. Therefore, before we can draw firm conclusions about the ATU abilities within and especially between species, further studies need to be conducted that allow more direct comparisons. Future research could adapt tasks that have been used with non-human animals to humans (and other species), and tasks used in the current study could be adapted for use with non-human animals.

Evidence for whether there was a cognitive hierarchy of the three examined ATU types was inconclusive. Experiment 1 showed that Tool set use yielded higher ATU and Correct success rates than Sequential tool use, as well as higher ATU rates than Metatool use, suggesting tentative evidence for the existence of such a hierarchy. However, there might be an alternative hypothesis explaining the differences between Tool set and Sequential tool use: the low ATU and Correct success rates in the Sequential tool use rates could be explained by the fact that both tasks consisted of two separate apparatuses, which might have been an additional source of difficulty for the children. Yet there is no evidence in the literature that the use of (spatially separate) apparatuses or platforms poses an additional cognitive demand in tool use tasks for non-human animals or children (Gruber et al., 2019; Jackson, 1942; Martin-Ordas et al., 2012; Miller et al., 2020; Mulcahy et al., 2005; Warden et al., 1940). Experiment 2 controlled for the number of apparatuses (using a single apparatus for both ATU versions) and found that Tool set and Sequential tool use did not differ in either ATU nor Correct success rates. However, performance in Experiment 2 was found to be substantially affected by task-specific features unrelated to ATU type. For example, in the Tool set use task of the box apparatus, some children failed to engage in ATU because they did not use enough force to pierce the barrier or tried to circumvent it – which might have appeared as a potentially successful strategy as the distance between the tool and the reward was reduced and this might have made it difficult for children to recognise this as a wrong attempt. Similarly, in the Tool set use task of the *tube* apparatus (i.e. same ATU type but different apparatus) many children failed as they accidentally dropped the stick into the apparatus and were unable to retrieve it. These inadvertent design features might have artificially decreased children's Tool set use abilities, and thus hindered the detection of a potential cognitive hierarchy. This emphasises the importance of investigating tool use with a multitude of tasks sharing the cognitive demand in question to be able to describe the emergence and development of these skills abstracted from task idiosyncrasies (Völter et al., 2018).

While the question whether there is a cognitive hierarchy among the three investigated ATU types requires further research, attempts have been made to create a hierarchy on a broader level, sorting tool use and other tool-related behaviours by their suggested cognitive complexity (Neldner, 2020; Putt et al., 2022; Visalberghi & Fragaszy, 2012). While ATU was not explicitly included in Neldner (2020), we suggest that it could be added at the upper end of the 'simple tool use' category to represent an intriguing link to more cognitively challenging tool-related behaviours such as tool innovation (Figure 4).

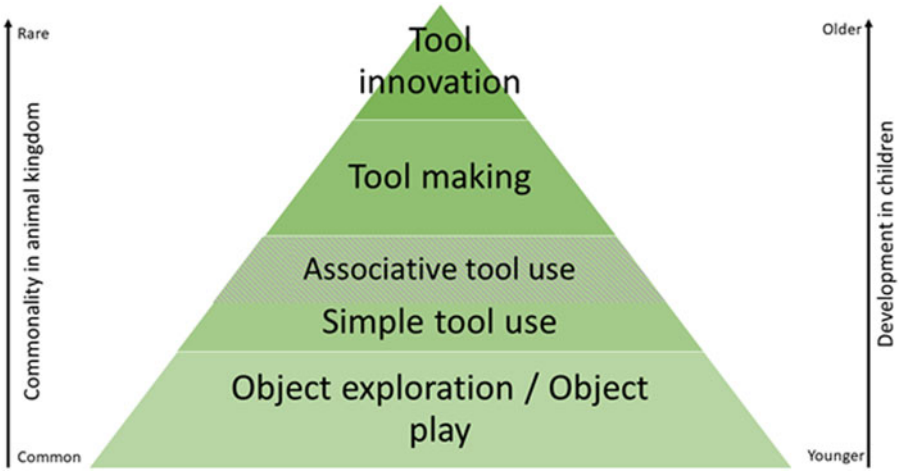

Figure 4. Adapted from Neldner's (2020; Figure 1.1) proposed hierarchy of tool-related behaviours. Here, we explicitly added Associative tool use at the upper end of the Simple tool use category.

Like simple tool use, ATU 'requires goal-directed, relational action between multiple objects' (Neldner, 2020, p. 16), but involves a greater number of objects, posing greater demands on working memory, inhibitory control, causal reasoning and imagination. In contrast to tool innovation, ATU requires no new tools to be imagined; however, it requires imagining the correct sequence of actions and how the objects involved relate to each other. Lastly, ATU is an intriguing addition to such a hierarchy, as it consists of one sub-group of behaviours which links the tool *use* and tool *making* categories, namely *Secondary tool use*.

In Secondary tool use, a tool is used to manufacture (or structurally modify) another tool (Shumaker et al., 2011). It seems to be unique to humans as it has not been reported in any wild non-human animals, and unenculturated, captive chimpanzees have shown no evidence for spontaneous Secondary tool use (Bandini et al., 2021). Although two bonobos and one orangutan were able to make and subsequently use stone tools, this was only after extensive periods of teaching, training and sometimes moulding by humans (Roffman et al., 2012; Schick et al., 1999; Toth et al., 1993; Toth & Schick, 2009; Wright, 1972; for a review see Bandini et al., 2021). Moreover, even after years of practice, the apes did not overcome certain cognitive (and morphological) restrictions to produce tools similar to the earliest hominin stone tools (Toth & Schick, 2009). Secondary tool use is regarded as a major cognitive and technological breakthrough in human evolution which has itself shaped human cognition and culture substantially owing to the coevolution of technological advances and cognitive capacities, such as working memory, planning and technical reasoning (Haidle, 2010; Lombard & Haidle, 2012; Osiurak, 2020; Read, 2008). Future studies could create Secondary tool use latent solution tests for children to find out more about the development of this ATU type.

Tool set and Sequential tool use were more difficult than Multifunctional tool use. This could be because of the larger relational complexity of ATU as, by definition, more objects are involved. This would imply a greater demand on executive function. As inhibitory control, working memory and attention shifting are themselves still developing during childhood and beyond (Best et al., 2009, 2011; Garon et al., 2008), executive function could indeed be a critical developmental bottleneck contributing to the difficulty of ATU for children (for a similar discussion of the importance of working memory for the *evolution* of ATU see e.g. Haidle, 2010; Read, 2017; Read et al., 2021; Wynn & Coolidge, 2014). Indeed, there is evidence suggesting that toddlers' performance in a tool selection task is correlated with their performance in an inhibition and a short-term memory task (Pauen & Bechtel-Kuehne, 2016) and that preschoolers' ability to make tools after observing a demonstrator

can be predicted by their score on a response inhibition task (S. R. Beck et al., 2016; Gönül et al., 2018). More studies are needed to compare children's performance in ATU and executive function tasks to better understand the cognitive demands involved in ATU and problem-solving in general.

In non-humans, ATU tasks have been used in latent solution tests as well as to study problem-solving and causal cognition more generally. We hope that the tasks introduced here will be a valuable addition to the pool of tests for investigating these and other topics in children and that they will be used and amended in future research.

Acknowledgements. We thank all nurseries, Thinktank Birmingham science museum, parents and children for their help and participation, Alice Coombes and Charlotte Wilks for assistance with data collection and for reliability coding, Lauren Cooper for assistance with data collection, Nuria Melisa Morales García (https://www.sciencegraphicdesign.com/) for the drawings of the apparatuses, and the editor and two anonymous reviewers for their very helpful comments. Ethical approval was granted by the University of Birmingham, UK, STEM Ethical Review Committee.

Author contributions. CT acquired the funding; ER, CT, SRB and IAA designed the study; ER collected, analysed and visualised the data; ER, CT, SRB and IAA discussed the results; ER and ZL conducted the literature review; ER, CT and SRB wrote the article; IAA and ZL edited and commented on the article; CT, SRB and IAA provided supervision.

Financial support. This work was supported by a PhD studentship of the College of Life and Environmental Sciences, University of Birmingham, UK (ER), and two Economic and Social Research Council Grants (CT, grant number ES/K008625/1; IAA and SRB, grant number ES/J023485/1). The funding sources had no involvement in study design, collection, analysis and interpretation of the data, or writing the paper, or in the decision to submit the article for publication. Access to some of the not publicly available studies listed in the literature review was funded by ER, and commissioning of the drawings of the apparatuses was funded by ER and CT (carried out after the end of formal funding).

Conflicts of interest. The authors declare no competing interests.

Research transparency and reproducibility. The data and script for both experiments presented in this study and Table S1 can be found on the OSF website: https://osf.io/d3pz5/?view_only=ba368e675f324a56a7cad600d6c39581

Supplementary material. To view supplementary material for this article, please visit https://doi.org/10.1017/ehs.2022.4

References

Alpert, A. (1928). *The solving of problem-situations by preschool children. An analysis* (PhD thesis). Teachers College, Columbia University, New York City.

Anderson, J. R., & Henneman, M. C. (1994). Solutions to a tool-use problem in a pair of Cebus apella. *Mammalia*, *58*(3), 351–362. https://doi.org/10.1515/mamm.1994.58.3.351

Baayen, R. H. (2008). *Analyzing linguistic data. A practical introduction to statistics using R*. Cambridge University Press.

Bandini, E., Motes-Rodrigo, A., Archer, W., Minchin, T., Axelsen, H., Hernandez-Aguilar, R. A., … Tennie, C. (2021). Naïve, unenculturated chimpanzees fail to make and use flaked stone tools [version 1; peer review: 3 approved]. *Open Res Europe*, *1*(20), 1–19. https://doi.org/10.12688/openreseurope.13186.1

Bandini, E., & Tennie, C. (2017). Spontaneous reoccurrence of 'scooping', a wild tool-use behaviour, in naïve chimpanzees. *PeerJ*, *5*, e3814. https://doi.org/10.7717/peerj.3814

Bandini, E., & Tennie, C. (2020). Exploring the role of individual learning in animal tool-use. *PeerJ*, *8*, e9877. https://doi.org/10.7717/peerj.9877

Barrett, T. M., Davis, E. F., & Needham, A. (2007). Learning about tools in infancy. *Developmental Psychology*, *43*(2), 352–368. https://doi.org/10.1037/0012-1649.43.2.352

Bates, D. M., Maechler, M., Bolker, B., & Walker, S. (2013). *lme4: Linear mixed-effects models using Eigen and S4*.

Bates, E., Carlson-Luden, V., & Bretherton, I. (1980). Perceptual aspects of tool using in infancy. *Infant Behavior and Development*, *3*(1), 127–140. https://doi.org/10.1016/S0163-6383(80)80017-8

Bechtel, S., Jeschonek, S., & Pauen, S. (2013). How 24-month-olds form and transfer knowledge about tools: The role of perceptual, functional, causal, and feedback information. *Journal of Experimental Child Psychology*, *115*(1), 163–179. https://doi.org/10.1016/j.jecp.2012.12.004

Beck, S. R., Apperly, I. A., Chappell, J., Guthrie, C., & Cutting, N. (2011). Making tools isn't child's play. *Cognition*, *119*, 301–306. https://doi.org/10.1016/j.cognition.2011.01.003

Beck, S. R., Williams, C., Cutting, N., Apperly, I. A., & Chappell, J. (2016). Individual differences in children's innovative problem-solving are not predicted by divergent thinking or executive functions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1690), 20150190. https://doi.org/10.1098/rstb.2015.0190

Bernstein-Kurtycz, L. M., Hopper, L. M., Ross, S. R., & Tennie, C. (2020). Zoo-housed chimpanzees can spontaneously use tool sets but perseverate on previously successful tool-use methods. *Animal Behavior and Cognition*, *7*(3), 288–309.

Best, J. R., Miller, P. H., & Jones, L. L. (2009). Executive functions after age 5: Changes and correlates. *Developmental Review*, *29*(3), 180–200. https://doi.org/10.1016/j.dr.2009.05.002

Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and Individual Differences*, *21*(4), 327–336. https://doi.org/10.1016/j.lindif.2011.01.007.

Bjorklund, D. F., & Gardiner, A. K. (2012). Object play and tool use: Developmental and evolutionary perspectives. *The Oxford handbook of the development of play*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195393002.013.0013

Boesch, C. (2013). Ecology and cognition of tool use in chimpanzees. In *Tool use in animals. Cognition and ecology* (pp. 21–47). Cambridge University Press.

Bolwig, N. (1963). Observations on the mental and manipulative abilities of a captive baboon (*Papio doguera*). *Behaviour*, *22*(1), 24–40.

Breyel, S., & Pauen, S. (2021). The beginnings of tool innovation in human ontogeny: How three- to five-year-olds solve the vertical and horizontal tube task. *Cognitive Development*, *58*, 101049. https://doi.org/10.1016/j.cogdev.2021.101049

Brown, A. L. (1990). Domain-specific principles affecting learning and transfer in children. *Cognitive Science*, *14*, 107–133. https://doi.org/10.1207/s15516709cog1401_6

Buskell, A., & Tennie, C. (accepted). Mere recurrence and cumulative culture at the margins. *British Journal for the Philosophy of Science*. https://doi.org/10.1086/717776

Carvalho, S., Cunha, E., Sousa, C., & Matsuzawa, T. (2008). Chaînes opératoires and resource-exploitation strategies in chimpanzee (Pan troglodytes) nut cracking. *Journal of Human Evolution*, *55*, 148–163. https://doi.org/10.1016/j.jhevol.2008.02.005

Chappell, J., Cutting, N., Apperly, I. A., & Beck, S. R. (2013). The development of tool manufacture in humans: what helps young children make innovative tools? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *368*(1630), 20120409. https://doi.org/10.1098/rstb.2012.0409

Chen, Z., Siegler, R. S., & Daehler, M. W. (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*(2), i–vii, 1–96. https://doi.org/10.1111/1540-5834.00072

Colbourne, J. A. D., Auersperg, A. M. I., Lambert, M. L., Huber, L., & Völter, C. J. (2021). Extending the reach of tooling theory: A neurocognitive and phylogenetic perspective. *Topics in Cognitive Science*, 1–25. https://doi.org/10.1111/tops.12554

Deák, G. O. (2014). Development of adaptive tool-use in early childhood. Sensorimotor, social, and conceptual factors. *Advances in Child Development and Behavior*, *46*, 149–181. https://doi.org/10.1016/B978-0-12-800285-8.00006-6

Dobson, A. J. (2002). *An introduction to generalized linear models*. Boca Raton: Chapman Hall/CRC.

Esseily, R., Rat-Fischer, L., Somogyi, E., O'Regan, K. J., & Fagard, J. (2016). Humour production may enhance observational learning of a new tool-use action in 18-month-old infants. *Cognition and Emotion*, *30*(4), 817–825. https://doi.org/10.1080/02699931.2015.1036840

Fagard, J., Rat-Fischer, L., & Kevin O'Regan, J. (2014). The emergence of use of a rake-like tool: A longitudinal study in human infants. *Frontiers in Psychology*, *5*(MAY), 1–12. https://doi.org/10.3389/fpsyg.2014.00491

Forstmeier, W., & Schielzeth, H. (2011). Cryptic multiple hypotheses testing in linear models: Overestimated effect sizes and the winner's curse. *Behavioral Ecology and Sociobiology*, *65*(1), 47–55. https://doi.org/10.1007/s00265-010-1038-5

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin*, *134*(1), 31–60. https://doi.org/10.1037/0033-2909.134.1.31

Gönül, G., Hohenberger, A., Corballis, M., & Henderson, A. M. E. (2019). Joint and individual tool making in preschoolers: From social to cognitive processes. *Social Development*, 1–17. https://doi.org/10.1111/sode.12373.

Gönül, G., Takmaz, E. K., Hohenberger, A., & Corballis, M. (2018). The cognitive ontogeny of tool making in children: The role of inhibition and hierarchical structuring. *Journal of Experimental Child Psychology*, *173*, 222–238. https://doi.org/10.1016/j.jecp.2018.03.017

Greif, M. L., & Needham, A. (2011). The development of human tool use early in life. In T. McCormack, C. Hoerl, & S. Butterfill (Eds.), *Tool use and causal cognition* (pp. 1–23). Oxford University Press. https://doi.org/10.1093/acprof

Gruber, R., Schiestl, M., Boeckle, M., Frohnwieser, A., Miller, R., Gray, R. D., … Taylor, A. H. (2019). New Caledonian crows use mental representations to solve metatool problems. *Current Biology*, *29*(4), 686-692.e3. https://doi.org/10.1016/j.cub.2019.01.008

Gurven, M., Kaplan, H., & Gutierrez, M. (2006). How long does it take to become a proficient hunter? Implications for the evolution of extended development and long life span. *Journal of Human Evolution*, *51*(5), 454–470. https://doi.org/10.1016/j.jhevol.2006.05.003

Haidle, M. N. (2010). Working-memory capacity and the evolution of modern cognitive potential. Implications from animal and early human tool use. *Current Anthropology*, *51*(1), S149–S166. https://doi.org/10.1086/650295

Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences*, *21*, 803–864.

Henrich, J. (2015). *The secret of our success. How culture is driving human evolution, domesticating our species, and making us smarter*. Princeton University Press.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*, 61–135. https://doi.org/10.1017/S0140525X0999152X

Hunt, G. R., Gray, R. D., & Taylor, A. H. (2013). Why is tool use rare in animals? In C. M. Sanz & C. Boesch (Eds.), *Tool use in animals. Cognition and ecology* (pp. 89–118). Cambridge University Press.

Jackson, T. A. (1942). Use of the stick as a tool by young chimpanzees. *Journal of Comparative Psychology*, *34*, 223–235.

Kahrs, B. A., Jung, W. P., & Lockman, J. J. (2013). When does tool use become distinctively human? Hammering in young children. *Child Development*, *85*(3), 1050–1061. https://doi.org/10.1111/cdev.12179

Kaplan, H. S., & Robson, A. J. (2002). The emergence of humans: The coevolution of intelligence and longevity with intergenerational transfers. *Proceedings of the National Academy of Sciences*, *99*(15), 10221–10226. https://doi.org/10.1073/pnas

Keen, R. (2010). The development of problem solving in young children: A critical cognitive skill. *Annual Review of Psychology*, *62*(1), 1–21. https://doi.org/10.1146/annurev.psych.031809.130730

Köhler, W. (1921). *Intelligenzprüfungen an Menschenaffen*. Springer.

Lancy, D. F. (2016). Playing with knives: The socialization of self-initiated learners. *Child Development*, *87*(3), 654–665. https://doi.org/10.1111/cdev.12498

Lancy, D. F. (2017). Homo Faber Juvenalis: a multidisciplinary survey of children as tool makers/users. *Childhood in the Past*, *10*(1), 72–90. https://doi.org/10.1080/17585716.2017.1316010

Lehner, S. R., Burkart, J. M., & van Schaik, C. P. (2011). Can captive orangutans (*Pongo pygmaeus abelii*) be coaxed into cumulative build-up of techniques? *Journal of Comparative Psychology*, *125*(4), 446–455. https://doi.org/10.1037/a0024413

Lethmate, J. (1982). Tool-using skills of orang-utans. *Journal of Human Evolution*, *11*, 49–64.

Lew-Levy, S., Kissler, S. M., Boyette, A. H., Crittenden, A. N., Mabulla, I. A., & Hewlett, B. S. (2020). Who teaches children to forage? Exploring the primacy of child-to-child teaching among Hadza and BaYaka Hunter-Gatherers of Tanzania and Congo. *Evolution and Human Behavior*, *41*(1), 12–22. https://doi.org/10.1016/j.evolhumbehav.2019.07.003

Lew-Levy, S., Pope, S. M., Haun, D. B. M., Kline, M. A., & Broesch, T. (2021). Out of the empirical box: A mixed-methods study of tool innovation among Congolese BaYaka forager and Bondongo fisher–farmer children. *Journal of Experimental Child Psychology*, *211*(105223). https://doi.org/10.1016/j.jecp.2021.105223

Lew-Levy, S., Reckin, R., Lavi, N., Cristóbal-Azkarate, J., & Ellis-Davies, K. (2017). How do hunter–gatherer children learn subsistence skills?: A meta-ethnographic review. *Human Nature*, *28*(4), 367–394. https://doi.org/10.1007/s12110-017-9302-2

Lockman, J. J. (2000). A perception–action perspective on tool use development. *Child Development*, *71*(1), 137–144. https://doi.org/10.1111/1467-8624.00127

Lombard, M., & Haidle, M. N. (2012). Thinking a bow-and-arrow set: Cognitive implications of Middle Stone Age bow and stone-tipped arrow technology. *Cambridge Archaeological Journal*, *22*, 237–264. https://doi.org/10.1017/S095977431200025X

Luria, A. R., & Vygotsky, L. (1930). *Ape, primitive man, and child: Essays in the history of behaviour*. Harvester Wheatsheaf.

Mannu, M., & Ottoni, E. B. (2009). The enhanced tool-kit of two groups of wild bearded capuchin monkeys in the Caatinga: Tool making, associative use, and secondary tools. *American Journal of Primatology*, *71*, 242–251. https://doi.org/10.1002/ajp.20642

Martin-Ordas, G., Schumacher, L., & Call, J. (2012). Sequential tool use in great apes. *PLoS ONE*, *7*(12). https://doi.org/10.1371/journal.pone.0052074

Matheson, E. (1931). A study of problem solving behavior in pre-school children. *Child Development*, *2*(4), 242–262. https://doi.org/10.2307/1125650

Matsuzawa, T. (1991). Nesting cups and metatools in chimpanzees. *Behavioral and Brain Sciences*, *14*, 570–571.

McCarty, M. E., Clifton, R. K., & Collard, R. R. (1999). Problem solving in infancy: The emergence of an action plan. *Developmental Psychology*, *35*(4), 1091–1101. https://doi.org/10.1037/0012-1649.35.4.1091

McCarty, M. E., Clifton, R. K., & Collard, R. R. (2001). The beginnings of tool use by infants and toddlers. *Infancy*, *2*(2), 233–256. https://doi.org/10.1207/S15327078IN0202_8

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models*. London: Chapman and Hall.

Metevier, C. M. (2006). *Tool-using in rhesus monkeys and 36-month-old children: Acquisition, comprehension, and individual differences* (PhD thesis). University of Massachusetts Amherst.

Miller, R., Gruber, R., Frohnwieser, A., Schiestl, M., Jelbert, S. A., Gray, R. D., … Clayton, N. S. (2020). Decision-making flexibility in New Caledonian crows, young children and adult humans in a multi-dimensional tool-use task. *PLoS ONE*, *15*(3), 1–20. https://doi.org/10.1371/journal.pone.0219874

Motes-Rodrigo, A., & Tennie, C. (2021). The method of local restriction: In search of potential great ape culture-dependent forms. *Biological Reviews*, *96*(4), 1441–1461. https://doi.org/10.1111/brv.12710

Mulcahy, N. J., Call, J., & Dunbar, R. I. (2005). Gorillas (*Gorilla gorilla*) and orangutans (*Pongo pygmaeus*) encode relevant problem features in a tool-using task. *Journal of Comparative Psychology*, *119*(1), 23–32. https://doi.org/10.1037/0735-7036.119.1.23

Nagell, K., Olguin, R. S., & Tomasello, M. (1993). Processes of social learning in the tool use of chimpanzees (*Pan troglodytes*) and human children (*Homo sapiens*). *Journal of Comparative Psychology* (*Washington, D.C. :* 1983), *107*(2), 174–186. https://doi.org/10.1037//0735-7036.107.2.174

Neldner, K. (2020). *The development and evolution of tool innovation in human children*. PhD thesis. University of Queensland.

Neldner, K., Reindl, E., Tennie, C., Grant, J., Tomaselli, K., & Nielsen, M. (2020). A cross-cultural investigation of young children' s spontaneous invention of tool use behaviours. *Royal Society Open Science*, *7*(192240). https://doi.org/10.1098/rspb.2015.2402

Nielsen, M., Cucchiaro, J., & Mohamedally, J. (2012). When the transmission of culture is child's play. *PLoS ONE*, *7*(3). https://doi.org/10.1371/journal.pone.0034066

Nielsen, M., Tomaselli, K., Mushin, I., & Whiten, A. (2014). Exploring tool innovation: A comparison of Western and Bushman children. *Journal of Experimental Child Psychology*, *126*, 384–394. https://doi.org/10.1016/j.jecp.2014.05.008

Nieuwenhuis, R., te Grotenhuis, M., & Pelzer, B. (2012). influence.ME: Tools for detecting influential data in mixed effects models. *The R Journal*, *4*(2), 38–47.

O'Hara, M., Mioduszewska, B., Mundry, R., Yohanna, Haryoko, T., Rachmatika, R., … Auersperg, A. M. I. (2021). Wild Goffin's cockatoos flexibly manufacture and use tool sets. *Current Biology*, *31*, 1–9. https://doi.org/10.1016/j.cub.2021.08.009

Osiurak, F. (2020). *The tool instinct*. Wiley. https://doi.org/10.1002/9781119720348

Pauen, S., & Bechtel-Kuehne, S. (2016). How toddlers acquire and transfer tool knowledge: Developmental changes and the role of executive functions. *Child Development*, *87*(4), 1233–1249. https://doi.org/10.1111/cdev.12532

Piaget, J. (1952). *The origins of intelligence in children*. International University Press.

Putt, S. S. J., Anwarzai, Z., Holden, C., Ruck, L., & Schoenemann, P. T. (2022). The evolution of combinatoriality and compositionality in hominid tool use: a comparative perspective. *International Journal of Primatology*. https://doi.org/10.1007/s10764-021-00267-7

Rat-Fischer, L., O'Regan, J. K., & Fagard, J. (2012). The emergence of tool use during the second year of life. *Journal of Experimental Child Psychology*, *113*(3), 440–446. https://doi.org/10.1016/j.jecp.2012.06.001

Rawlings, B. S., Flynn, E. G., & Kendal, R. L. (2021). Personality predicts innovation and social learning in children: Implications for cultural evolution. *Developmental Science*, *e*13153. https://doi.org/10.1111/desc.13153

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Read, D. W. (2008). Working memory: A cognitive limit to non-human primate recursive thinking prior to hominid evolution. *Evolutionary Psychology*, *6*(4), 676–714.

Read, D. W. (2017). Quantitative differences between the working memory of chimpanzees and humans give rise to qualitative differences: Subitizing and cranial development. *UCLA: Human Complex Systems*., Retrieved from https://escholarship.org/uc/item/5d.

Read, D. W., Manrique, H. M., & Walker, M. J. (2021). On the working memory of humans and great apes: Strikingly similar or remarkably different? *Neuroscience & Biobehavioral Reviews*. https://doi.org/10.1016/J.NEUBIOREV.2021.12.019

Reindl, E., Bandini, E., & Tennie, C. (2018). The zone of latent solutions and its relation to the classics: Vygotsky and Köhler. In L. D. di Paolo, F. di Vicenzo, & F. de Petrillo (Eds.), *Evolution of primate social cognition* (pp. 231–248). Springer.

Reindl, E., Beck, S. R., Apperly, I. A., & Tennie, C. (2016). Young children spontaneously invent wild great apes' tool-use behaviours. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1825), 20152402. https://doi.org/10.1098/rspb.2015.2402

Reindl, E., Parkash, D., Völter, C. J., & Seed, A. M. (2021). Thinking inside the box: Mental manipulation of working memory contents in 3- to 7-year-old children. *Cognitive Development*, *59*(101068). https://doi.org/10.1016/j.cogdev.2021.101068

Reindl, E., & Wronski, C. (2022). The role of social context for innovation in children. Preregistered study, February 14. Retrieved from osf.io/rvs4y

Roffman, I., Savage-Rumbaugh, S., Rubert-Pugh, E., Ronen, A., & Nevo, E. (2012). Stone tool production and utilization by bonobo-chimpanzees (*Pan paniscus*). *Proceedings of the National Academy of Sciences of the United States of America*, *109*(36), 14500–14503. https://doi.org/10.1073/pnas.1212855109

Schick, K. D., Toth, N., Garufi, G., Savage-Rumbaugh, E. S., Rumbaugh, D., & Sevcik, R. (1999). Continuing investigations into the stone tool-making and tool-using capabilities of a bonobo (*Pan paniscus*). *Journal of Archaeological Science*, *26*, 821–832.

Shumaker, R. W., Walkup, K. R., & Beck, B. B. (2011). *Animal tool behavior. The use and manufacture of tools by animals*. Johns Hopkins University Press.

Somogyi, E., Ara, C., Gianni, E., Rat-Fischer, L., Fattori, P., O'Regan, J. K., & Fagard, J. (2015). The roles of observation and manipulation in learning to use a tool. *Cognitive Development*, *35*, 186–200. https://doi.org/10.1016/j.cogdev.2015.06.002

Stewart, C., & Disotell, T. R. (1998). Primate evolution – In and out of Africa. *Current Biology*, *8*, 582–588.

Sugiyama, Y. (1997). Social tradition and the use of tool-composites by wild chimpanzees. *Evolutionary Anthropology*, *6*, 23–27.

Taylor, A. H., Elliffe, D., Hunt, G. R., & Gray, R. D. (2010). Complex cognition and behavioural innovation in New Caledonian crows. *Proceedings. Biological Sciences/The Royal Society*, *277*(1694), 2637–2643. https://doi.org/10.1098/rspb.2010.0285

Taylor, A. H., Elliffe, D. M., Hunt, G. R., Emery, N. J., Clayton, N. S., & Gray, R. D. (2011). New Caledonian crows learn the functional properties of novel tool types. *PLoS ONE*, *6*(12). https://doi.org/10.1371/journal.pone.0026887

Taylor, A. H., Hunt, G. R., Holzhaider, J. C., & Gray, R. D. (2007). Spontaneous metatool use by New Caledonian crows. *Current Biology*, *17*, 1504–1507. https://doi.org/10.1016/j.cub.2007.07.057

Tennie, C., Bandini, E., van Schaik, C. P., & Hopper, L. M. (2020). The zone of latent solutions and its relevance to understanding ape cultures. *Biology and Philosophy*, *35*(5), 1–42. https://doi.org/10.1007/s10539-020-09769-9

Tennie, C., Braun, D. R., Premo, L. S., & McPherron, S. P. (2016). The island test for cumulative culture in the Paleolithic. In *The nature of culture* (pp. 121–133). Springer. https://doi.org/10.1007/978-94-017-7426-0

Tennie, C., Call, J., & Tomasello, M. (2009). Ratcheting up the ratchet: On the evolution of cumulative culture. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1528), 2405–2415. https://doi.org/10.1098/rstb.2009.0052

Toth, N., & Schick, K. (2009). The Oldowan: The tool making of early hominins and chimpanzees compared. *Annual Review of Anthropology*, *38*, 289–305. https://doi.org/10.1146/annurev-anthro-091908-164521

Toth, N., Schick, K. D., Savage-Rumbaugh, E. S., Sevcik, R. A., & Rumbaugh, D. M. (1993). Pan the tool-maker: Investigations into the stone tool-making and tool-using capabilities of a bonobo (*Pan paniscus*). *Journal of Archaeological*, *20*, 81–91.

van Schaik, C. P., Ancrenaz, M., Borgen, G., Galdikas, B., Knott, Cheryl, D., … Merrill, M. (2003). Orangutan cultures and the comparative study of culture. *Science*, *299*, 102–105. https://doi.org/10.1126/science.1078004

van Schaik, C. P., Ancrenaz, M., Djojoasmoro, R., Knott, C. D., Morrogh-Bernard, H. C., Odom, K., … van Noordwijk, M. A. (2009). Orangutan cultures revisited. In *Orangutans: Geographic variation in behavioral ecology and conservation* (pp. 299–309).

van Schaik, C. P., Deaner, R. O., & Merrill, M. Y. (1999). The conditions for tool use in primates: Implications for the evolution of material culture. *Journal of Human Evolution*, *36*(6), 23. https://doi.org/S0047-2484(99)90304-4 [pii]\r10.1006/jhev.1999.0304

Visalberghi, E., & Fragaszy, D. (2012). What is challenging about tool use? The capuchin's perspective. In T. R. Zentall & E. A. Wasserman (Eds.), *Comparative cognition: Experimental explorations of animal intelligence* (pp. 777–799). Oxford University Press.

Voigt, B., Pauen, S., & Bechtel-Kuehne, S. (2019). Getting the mouse out of the box: Tool innovation in preschoolers. *Journal of Experimental Child Psychology*, *184*, 65–81. https://doi.org/10.1016/j.jecp.2019.03.005

Völter, C. J., Tinklenberg, B., Call, J., & Seed, A. M. (2018). Comparative psychometrics: Establishing what differs is central to understanding what evolves. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1756). https://doi.org/10.1098/rstb.2017.0283

Warden, C. J., Koch, A. M., & Fjeld, H. A. (1940). Instrumentation in cebus and rhesus monkeys. *The Journal of Genetic Psychology*, *56*, 297–310.

Westergaard, G. C., Greene, J. A., Menuhin-Hauser, C., & Suomi, S. J. (1996). The use of naturally-occurring copper and iron tools by monkeys: Possible implications for the emergence of metal-tool technology in hominids. *Human Evolution*, *11*(1), 17–25.

Westergaard, G. C., Lundquist, a. L., Kuhn, H. E., & Suomi, S. J. (1997). Ant-gathering with tools by captive tufted capuchins (*Cebus apella*). *International Journal of Primatology*, *18*(1), 95–103. https://doi.org/10.1023/A:1026345307953

Westergaard, G. C., & Suomi, S. J. (1994a). A simple stone-tool technology in monkeys. *Journal of Human Evolution*, *27*, 399–404.

Westergaard, G. C., & Suomi, S. J. (1994b). Stone-tool bone-surface modification by monkeys. *Current Anthropology*, *35*(470), 468. http://www.journals.uchicago.edu/t-and-c

Whiten, A., Goodall, J., Mcgrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., … Boesch, C. (1999). Cultures in chimpanzees. *Nature*, *399*, 15–18.

Whiten, A., Goodall, J., Mcgrew, W. C., Nishida, T., Reynolds, V., Sugiyama, Y., … Boesch, C. (2001). Charting cultural variation in chimpanzees. *Behaviour*, *138*, 1481–1516.

Willatts, P. (1984). The stage-IV infant's solution of problems requiring the use of supports. *Infant Behavior and Development*, *7*(2), 125–134. https://doi.org/10.1016/S0163-6383(84)80053-3

Wimpenny, J. H., Weir, A. A. S., Clayton, L., Rutz, C., & Kacelnik, A. (2009). Cognitive processes associated with sequential tool use in New Caledonian crows. *PLoS ONE*, *4*(8). https://doi.org/10.1371/journal.pone.0006471

Wright, R. v. (1972). Imitative learning of a flaked tool technology – The case of an orangutan. *Mankind*, *8*, 296–306.

Wynn, T., & Coolidge, F. L. (2014). Technical cognition, working memory and creativity. *Pragmatics and Cognition*, *22*(1), 45–63. https://doi.org/10.1075/pc.22.1.03wyn

Cite this article: Reindl E, Tennie C, Apperly IA, Lugosi Z, Beck SR (2022). Young children spontaneously invent three different types of associative tool use behaviour. *Evolutionary Human Sciences* **4**, e5, 1–23. https://doi.org/10.1017/ehs.2022.4