The role of timing and prototypical causality on how preschoolers fast-map novel verb meanings

Kirsten Abbot-Smith[1], Mutsumi Imai[2], Samantha Durrant[3] and Erika Nurmsoo[1]

[1]School of Psychology, University of Kent, UK

[2]Keio University, Japan

[3]ESRC LuCiD Centre & School of Psychology, University of Liverpool, UK

Contact: Kirsten Abbot-Smith
Kent Child Development Unit
School of Psychology
University of Kent
Keynes College
Canterbury
CT2 7NP, UK
Tel: + 44 1227 823016
Email: K.Abbot-Smith@kent.ac.uk

Abstract

In controlled contexts, young children find it more difficult to learn novel words for actions than words for objects: Imai et al. (2008) found that English-speaking three-year-olds mistakenly choose a novel object as a referent for a novel verb about 42% of the time despite hearing the verb in a transitive sentence. The current two studies investigated whether English three- and five-year-old children would find resultative actions easier (since they are prototypically causative) than the non-resultative, durative event types used in Imai et al.'s studies. The reverse was true. Furthermore, if the novel verbs were taught on completion of the action, this did not improve performance, which contrasts with previous findings (e.g. Tomasello & Kruger, 1992). Our resultative actions were punctual, change-of-location events which may be less visually salient than the non-resulative, durative actions. Visual salience may play a greater role than does degree of action causality in the relative ease of verb learning even at three years. (*158 words*)

Keywords: *verb learning, pre-schoolers, pointing, resultative, transitive*

Introduction

Children learn a large number of names for objects at a surprisingly young age and often with surprisingly few exemplars (e.g. Fenson et al., 2004; see Childers & Tomasello, 2006, for a review). Words for actions, in contrast, are relatively difficult to learn (e.g. Genter, 1982, 2006; Childers & Tomasello, 2002; Imai, Haryu, & Okada, 2005). While there are cross-linguistic differences regarding whether nouns always outnumber verbs in early child spontaneous speech (e.g. Tardif, 1996), in experimental word-learning contexts even children learning languages with pervasive argument ellipsis appear to find it more difficult to learn words for actions than words for objects (e.g. Imai et al., 2008; see also Kim, McGregor & Thompson, 2000, for naturalistic speech) .

There are several possible reasons for this relative difficulty in learning words for actions. One possibility is that young children may be biased to map new words onto objects for which they do not have a name (see Markman, 1994; Markson & Bloom, 1997). This bias may push children to consider novel objects as candidates for the referent of a novel verb. For example, Kersten and Smith (2002) found that 3½ to 4-year-old children attended equally to novel objects and their actions when hearing a novel verb describing the object's path or manner of motion. (In fact, a number of studies have found that young preschool children often have difficulty generalising words for actions when other aspects of the original learning event are changed, such as

the agent, e.g. Maguire, Hirsh-Pasek, Golinkoff & Brandone, 2008). In a particularly strong demonstration, Imai and colleagues (Imai et al 2005, 2008) showed that while 3- and 5-year-old children were proficient at mapping a novel noun to a novel object, only 5-year-olds learned novel verbs appropriately. In these studies, in the Exposure phase for each item children were presented with videos of actors performing novel actions (e.g., twisting an object so that it bends in the middle) on novel objects (e.g., a rubber dog toy shaped like two triangles stuck together). During these videos, children heard a sentence that either used a novel noun (e.g., "Look! This is a *moop*!") or a novel verb (e.g., "Look! She is *mooping* it!"). In the Forced-Choice phase for each item, they were asked to find another instance of the novel word (e.g., "Where is the *moop*?" or "Where is she *mooping* it?") whereby they had to choose between two scenes: in the Same Action video, the actor performed the same action on a different object, and in the Same Object video, the actor performed a different action on the same object. Although 3- and 5-year-old children were both able to correctly choose at above chance levels the Same Object video (i.e. the target) on Noun trials, only five-year-old English-speaking children chose the correct Same Action scene (i.e. the target) more often than chance on Verb trials; English-speaking three-year-olds were at chance (Imai et al, 2008). Interestingly, these results were found even for children learning languages with pervasive argument-drop, namely Japanese and Mandarin.

Therefore, Imai et al.'s (2005, 2008) findings suggest that three (and sometimes five) year olds struggle to fast-map onto actions the meanings of novel verbs heard in transitive sentence frames. This stands in stark contrast to a body of research emerging from the syntactic bootstrapping literature (e.g. Scott & Fisher, 2009). Here, much younger children can fast-map novel verbs onto the correct novel event when asked to choose between two causative events, such as a) one where a duck is making a bunny rock by pulling its legs versus b) one where a bunny is spinning a duck around in a chair (e.g. Gertner, Fisher & Eisengart, 2006; Naigles, 1990; see also Noble, Rowland and Pine, 2011).

The key difference is that these latter studies did not contain novel objects; 'bunny' and 'duck' are words which are in the receptive vocabularies of one-year-olds and thus can be ruled out as potential referents for the novel verb (Gertner et al., 2006; Naigles, 1990; Noble et al., 2011). However, recently two looking-time habituation studies have found that even around 1½ years children can map novel words to actions and extend these to scenes with the same action but a different novel object during essentially one-trial learning (e.g. Chan et al, 2011; Oshima-Takane, Ariyama, Kobayashi, Katerlos & Poulin-Dubois, 2011). One crucial difference between the paradigms using the looking-time habituation studies, on the one hand, and the same-novel-action-vs.-same-novel-object paradigms used in Imai et al. (2005, 2008) and

Kerstin and Smith (2002) is that in the habituation paradigm children are never forced to choose between the same object versus the same action on a given trial.

That said, there are two studies by Arunachalam and Waxman (2011, 2015) in which two-year-olds did have to choose between the Same Object versus Same Action video clips on the test trial and indeed by pointing rather than merely via eye-gaze preference. Most crucially for our current purposes, however, in these studies the choice was not between a novel action and a novel object; rather the objects were familiar ones such as 'balloon'. Thus, the children in Arunchalam and Waxman's studies could have simply chosen the same action via mutual exclusivity (e.g. Markman, 1991), or an equivalent process, which is not an option for children tested in the same-novel-action-vs.-same-novel-object paradigm. Therefore, overall findings to date suggest that young children have difficulty fast-mapping novel verbs onto novel actions when novel objects are co-present (at least when social-pragmatic cues are removed, cf. Tomasello & Akhtar, 1995, for evidence that 27-month-olds can map correctly when given discourse novelty and / or intention-reading cues).

A second possibility for why verbs are more difficult to learn than nouns is that actions do not always have clear-cut beginning and end points (e.g. Golinkoff & Hirsh-Pasek, 2008; Gentner, 1982), so they may be more difficult to parse from the scene than objects. Transitive verbs in early child language frequently denote causative actions; that is, actions in which one entity (the agent) affects another (the patient) in some way.

Many causative actions may involve the agent moving the patient in some manner (e.g. spinning the patient around in a chair or flopping the patient iteratively up and down) but these 'non-resultative' causative actions do not have a clear end-point. A number of theorists have argued that prototypical causative actions are 'resultative', that is, they involve a result, typically a change of state (e.g. *wash*) or a change of location (e.g. *hurl*) (e.g. Gropen, Pinker & Hollander, 1991: 162; see also Slobin, 1981). Resultative actions end when the change has been made to the target object, so they may be easier to identify from the scene (see Brandone, Pence, Golinkoff & Hirsh-Pasek, 2007, for some suggestive evidence in this direction)

Importantly, the actions used by Imai et al. (2005, 2008) were not prototypically causative since they were not resultative because the object did not undergo a change of state or location. Rather, the actions were all of some duration involving iterative non-resultative actions on an object (e.g. the agent repeatedly rolling an object between her palms). As a result, we cannot determine whether children's difficulty in Imai et al. (2005, 2008)'s studies were due to the presence of the novel object, the type of action being taught or a combination of the two.

In addition, there is evidence that children's learning of a new verb is influenced by the timing of the presentation of the new label with respect to the action demonstration. Ambalu, Chiat and Pring (1997) found that children age 2;3 to 3;6 years learned novel verbs for non-resultative actions (e.g., spinning an object around) better if

the verbs were taught prior to the action. However, for a resultative event (e.g., stamping paper with a printing stamp), the verb was comprehended best when taught on completion of the action. Similarly, Tomasello and Kruger (1992) found that more 24-month-olds were 'comprehenders' (as opposed to non-comprehenders) of a novel verb if it had been taught either prior to or on completion of a resultative novel event in which patients were rolled down a slope into a new location. In contrast, when it had been taught during the action, there were an equal number of comprehenders and non-comprehenders at test. In Imai et al (2005, 2008), children heard the novel verbs concurrently with the demonstration. Thus teaching and testing a novel verb on completion of the action might improve performance.

The current studies

In two studies, we explored children's learning of action words to determine whether resultative actions are learned better than non-resultative, and whether the timing of the naming differentially impacts the learning of these types of actions (Study 2). For both Studies we followed Imai et al. (2008) in that all actions were performed on novel objects, allowing for a strict test of children's verb learning. Our procedure in the Non-Resultative condition closely followed that of Imai et al (2008), also using novel objects matched to those used by Imai and colleagues (2005, 2008) as listed in Appendix 1. The only difference in procedure to Imai et al. (2008) is that we also used the novel verbs in the future tense (e.g. '*She's gonna moop it!*') prior to each action

demonstration during the Exposure phase for each novel verb. This was because findings from Tomasello & Kruger (1992) and Ambalu et al. (1997) indicate that hearing a verb prior to an action may be an optimal attention-getter. In our Resultative condition we used punctual actions which either involved a change of location (e.g. head-butting an object onto the floor) or a change of position (e.g. flipping an object over). Punctual actions were chosen because causative actions tend to be expressed by transitive verbs and these tend to refer to punctual actions (e.g. Meints, 1999).

In Study 1 our main research question was whether resultative actions are easier than non-resultative actions. Imai et al. (2008) found (for non-resultative actions) that five-year-old English-speaking children performed above chance when the novel verb was heard in a transitive argument structure (e.g. '*She's blicking it!*) but they performed at chance when the verb was heard in isolation (e.g. '*Look! Blicking*!), whereas English-speaking three-year-olds were at chance in both conditions with no evidence that this experimental manipulation affected their performance at all. Since Imai et al.'s (2008) five-year-olds were not at ceiling (i.e. they selected the correct Same Action clip 70% of the time), we only tested five-year-olds for Study 1 in our first exploration of the role of Event Type in this paradigm. In Study 2, we maintained our two Event Type conditions (Resultatives vs. Non-resultatives) but also investigated, with both 3- and 5-year-olds, how the timing of the label affects verb learning.

Study 1

Stimuli verification pre-studies

Prior to running Studies 1 and 2 we also carried out a Salience Control pilot for both Event Types in order to control for whether young children found particular clips more visually salient in some way. For this Salience Control pilot we presented the Forced-Choice component of each experimental item trial (both conditions) and asked 13 three-year-olds '*where is she mooping it*?' without the children having first seen the corresponding Exposure clip. The children pointed at chance indicating that it was not the case that the children would point at the target clips for reasons of visual salience alone. In addition we also showed pictures of all novel objects to seven five-year-olds to ensure that they did not have a name for these objects

Our first study aimed to determine whether English-speaking five-year-olds would learn novel verbs for resultative actions better than for non-resultative actions.

Design

There were two between-participants Event Type conditions. In the Non-Resultative condition, actions were iterative, durative, non-resultative events replicated from Imai et al. (2005, 2008), such as repeatedly tapping an object against one's thigh (see Appendix 1). In the Resultative condition, actions were events in which the object changed

location (e.g. agent head-butts object onto floor) or position (e.g. agent flips object over, see Appendix 2).

Participants

The children were pre-assigned to one of two between-participants conditions (Resultative vs. Non-Resultative). We included 17 five-year-olds in the Non-Resultative (replication) condition (Mean age = 63.76 months, range 60-70 months, 47% boys) and 17 five-year-olds in the Resultative condition (Mean age = 63.94 months, range 60-69 months, 47% boys). The two groups did not differ in terms of age ($t(32) = 0.16$, $p = .872$, $d = -0.06$). All children were monolingual, typically-developing speakers of British English and we excluded children who scored more than 1 SD below the mean on the Expressive Vocabulary sub-test of the Clinical Evaluation of Language Fundamentals - Preschool (CELF-P, Semel, Wiig & Secord, 2004). Half in each condition were tested in the Kent Child Development Unit, and half in primary schools in Ashford, Kent (UK). The two conditions did not differ in CELF-P Expressive Vocabulary raw scores (Non-Resultative M = 28.65, range 19-34; Resultative M = 26.42, range = 16-34 $t(32) = 1.21$, $p = .24$, $d = 0.42$) whereby the possible maximum raw score was 40). In the Kent Child Development Unit, the parent sat directly behind the child, and in schools each child was tested individually in a quiet area.

Materials

Experimental item materials. As in Imai et al. (2005, 2008), each experimental item trial consisted of an Exposure phase clip followed by a pair of Forced-choice phase clips. Each Exposure phase clip showed a Caucasian woman carrying out a novel action (Exposure action) with a novel object (Exposure object). The Forced-choice clips each showed the same actor performing an action on an object. The Target clips were all Same-Action in which the agent carried out the Exposure action on a new novel object. The Foil clips were all Same-Object, in which the agent carried out a novel action on the Exposure object. Figure 1 illustrates the similarities between the foil (clip 7) vs. target clips (clip 8) and the Exposure clip (clips 1-6). In the Non-Resultative condition the Exposure actions were durative and iterative and we closely replicated both the actions and objects of the original studies by Imai et al. (2005, 2008) – see Appendix 1. In the Resultative condition the Exposure actions were punctual and had a lasting result; for example, the actor carried out an Olympic hammer-throw action in which she twirled an object above her head and then threw it so that it landed on the floor (see Appendix 2). The full list of experimental item trial actions and objects for each Forced-Choice phase is listed in the Appendices. For each phase of the experiment, the Resultative and Non-Resultative conditions were matched in length. To accommodate the fact that the resultative actions were of a relatively brief duration, we created non-resultative clips which were equal in length to those in the Non-Resultative conditions

and then looped the clips in both conditions. Thus, in the Forced-choice phase, for example, in both conditions, both the target and foil clips were looped five times (in synchrony) before freezing on the end still clip.

Warm-up phase materials. Each child first participated in a 'warm-up-phase' to ensure that children understood that the task was to point to the video clip out of two simultaneously running clips which matched what the Experimenter (E) said. In the warm-up phase, if children pointed to the wrong clip, they received corrective feedback. The warm-up phase was identical for all children and consisted of four trials, always in the same order. In the first warm-up trial, the children saw a clip of a woman eating a banana paired with a clip of a woman cutting a banana and the E asked '*Can you show me: where is she eating it*?*'*. The next trial showed kicking a ball paired with catching a ball and E asked '*Can you show me: where is she catching it*?*'*. Trials 3 and 4 of the warm-up phase were closely modelled on those in Imai et al. (2008), involved novel verbs and were parallel to the experimental item trials with the following the key difference; for these warm-up trials, the target clip of the Forced-choice phase showed not only the same novel action as the Exposure phase but also the same novel object. Further, for these warm-up novel trials, the foil differed from the Exposure phrase not only in terms of the novel action but also in terms of the novel object. (Thus, even if a child had a bias to map a novel word to the novel object, he or she should be correct for

the novel-verb warm-up trials). One trial out of warm-up trials three and four depicted a non-resultative action and the other involved a resultative action.

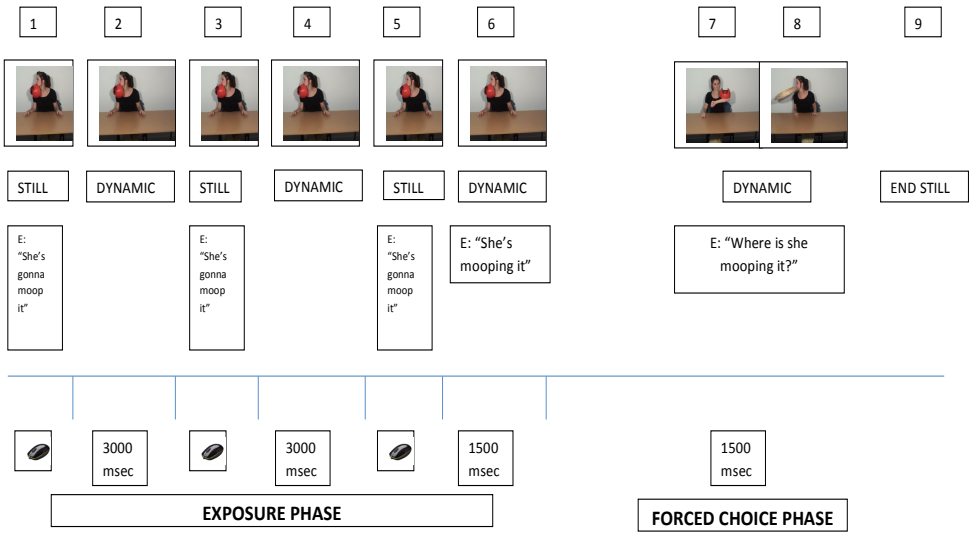
Procedure

For each trial, a 19" touch screen monitor recorded the children's responses, but E also noted down the child's selection. Each child was first administered the four warm-up trials, then the six experimental item trials and finally the CELF-P (Semel et al., 2004). The whole session lasted between 15-20 minutes. For the experimental item trials, the procedure is illustrated in Figure 1. Each experimental item trial consisted of two main phases: Exposure (learning), and forced-choice.

Exposure phase. Children first saw a still version of the clip (see clip 1, Fig 1) and heard E use the novel verb in a full transitive using the future 'gonna' (e.g. '*She's gonna moop it!*'). E then played the clip, which lasted three seconds (clip 2, Fig 1). This was repeated twice (see clips 3-6, Fig 1), and on the third repeat the Exposure video clip was observed for nine seconds (i.e. looped three times) and E used the same verb in a full transitive in the progressive tense three times (e.g. '*She is mooping it*!'). Thus, in each Exposure phase each child heard the novel verb used in a future tense form of the active transitive three times as well as in the progressive tense.

Forced-choice phase. In the subsequent Forced choice phase, the Target and Foil clips ran simultaneously for 15 seconds (i.e. looped five times) and children were asked '*where is she mooping it?*'. The location of the Target clip (left or right side) was counterbalanced both within and across participants, as was the order in which the actions and novel verbs were presented. Children's selection of clip (by pointing and/or vocalising) was recorded.

Figure 1: Timecourse of one representative test trial (tmecourse and timing identical for both Event Type conditions).



Data coding

Children scored a point each time they chose the Target clip on the Forced Choice trials. If a child pointed to both clips, that particular trial was removed from analysis (NA). This occurred for 2% of trials, If a child pointed to one clip but simultaneously said something which clearly indicated that he/she was NOT pointing to respond to the question (e.g. 'not that one'), data was coded according to what he/she said. (This only occurred on one or two occasions).

Results

Table 1: Mean (SDs) proportion correct points for experimental item trials in Study 1

Resultative	Non-resultative
.36 (.38)	.69 (.33)

In each condition there were six experimental item trials in total. To compare performance to chance, the dependent variable was the proportion of points to the correct (same-action) clip in the Forced-Choice phase. Participants in the Non-Resultative condition were significantly above chance at pointing correctly ($t(16) = 2.38$, $p < .05$), replicating Imai et al. (2008). In the Resultative condition, children did not point significantly above chance ($t(16) = 1.47$, $p = .16$).

To compare performance across our two experimental conditions, we used a binomial mixed effect model, which treats the dependent variable as a binary choice for each trial. Event Type was treated as a fixed effect and participants were treated as random effects, with random slopes for participants. The *p*-values were computed by comparing models with likelihood-ratio tests and chi-square values are reported. Participants were significantly more likely to correctly identify the Target clip in the Non-Resultative than Resultative condition ($b = 4.65$, SE = 2.03, $\chi^2(1) = 7.33$, $p < .01$). Contrary to our expectations, resultative actions appear to have been more difficult for our five-year-olds to learn than non-resultative actions.

Discussion

We expected that the resultative actions would help children parse the action from the scene, improving performance for resultative over non-resultative actions. However, the reverse was found, with non-resultative actions being better learned. It is possible that participants found the resultative actions more difficult because the visual brevity of punctual actions makes them more difficult to encode. This might be particularly problematic for young children, as their visual processing speed is slower than that of adults (e.g., Liss & Haith, 1970). While motion per se is known to be highly salient for infants who focus on and remember the details of actions such as bubble blowing versus hair brushing (e.g., Bahrick, Gogate & Ruiz, 2002), this may

only be the case for actions of lengthy duration. However, a counterargument to this could be that the key time-point of visual salience for resultatives is once the result (here: location change) is observable, which is after the action has occurred. That is, there may be in interaction between the timing of the linguistic model and the type of event the verb denotes, where durative, non-resultative events are best taught during the event while verbs denoting punctual, change-of-location events are best taught on completion of the event. This view receives some support from previous studies by Tomasello and Kruger (1992), Carr and Johnston (2001) and Ambalu et al. (1997).

Study 2

Study 1 did not provide support for the view that prototypically causative actions will be easier than non-prototypically causative actions to map to novel verbs heard in a transitive sentence frame. To explore the possibility that resultative actions (such as 'flip over' or 'volleyball-underhand-serve') would be learned better if these were taught and tested after the action had occurred, in Study 2 we presented the same stimuli as Study 1, but manipulated the timing (and tense) of the linguistic model. Some children heard the action descriptor once the action had occurred (in the past tense) and others heard it during the event (in the present tense). We also sought to explore younger children's abilities to learn resultative (such as 'flip over' or 'head-butt-away') vs. non-resultative actions (such as 'thigh-tap' or 'palm-roll').

Design and procedure

The design, materials and procedure were identical to Study 1 with the exception that we included 3-year-olds and we also manipulated the timing / tense of the verb differed. That is, Age Group (5-year-olds vs. 3-year-olds), Event Type (Resultative vs. Non-Resultative) and Timing (Ongoing vs. Past) were fully crossed between-subjects conditions. In the Past condition, the verb was modelled in the past tense after the Exposure clip had stopped (e.g. "*she mooped it")* and the question on the Forced-choice phases was similarly in the past tense after the clips had stopped (e.g. "*Show me: where did she moop it?*"). The Ongoing conditions were identical to those used in Study 1 (i.e. in each Exposure phase the verb was taught while the action was still ongoing as e.g. '*She is mooping it*' and in the Forced-choice phases tested as '*Show me: where is she mooping it*?).

Participants

All participants were tested in pre-schools and primary schools in southern England. As in Study 1, all were monolingual, typically-developing speakers of British English and we excluded children who scored more than 1 SD below the mean on the Expressive Vocabulary sub-test of the CELF-P (Semel et al. 2004). In Table 2 it can be seen that in Study 2 there were four between-subjects conditions for each age group (three-year-olds and five-year-olds). Table 2 also outlines the number of children included in each of

these between-subjects conditions for each age group, their mean age in months (and age range), the gender ratio, and their mean (and range) CELF Expressive Vocabulary scores.

Table 2 Demographics for Study 2
(NB: CELF expressive vocabulary possible maximum raw score = 40)

	Five-year-olds				**Three-year-olds**			
	Resultative		Non-Resultative (replication)		Resultative		Non-Resultative (replication)	
	Present	Past	Present	Past	Present	Past	Present	Past
STUDY 2 M age	63.84	63.35	63.5	64.33	42.76	43.18	42.65	42.53
Study 2 age range	60-71	60-70	60-69	60-70	36-47	39-47	37-47	40-46
Study 2 % boys	42%	47.8%	45%	19%	47%	47%	47%	47%
Study 2 M CELF vocab	25.58	28.17	28.3	29.57	17.88	16.41	15.94	16.88
Study 2 CELF range	16-34	17-37	20-38	22-40	10-25	11-22	10-30	10-22
Number	19	23	20	21	17	17	17	17

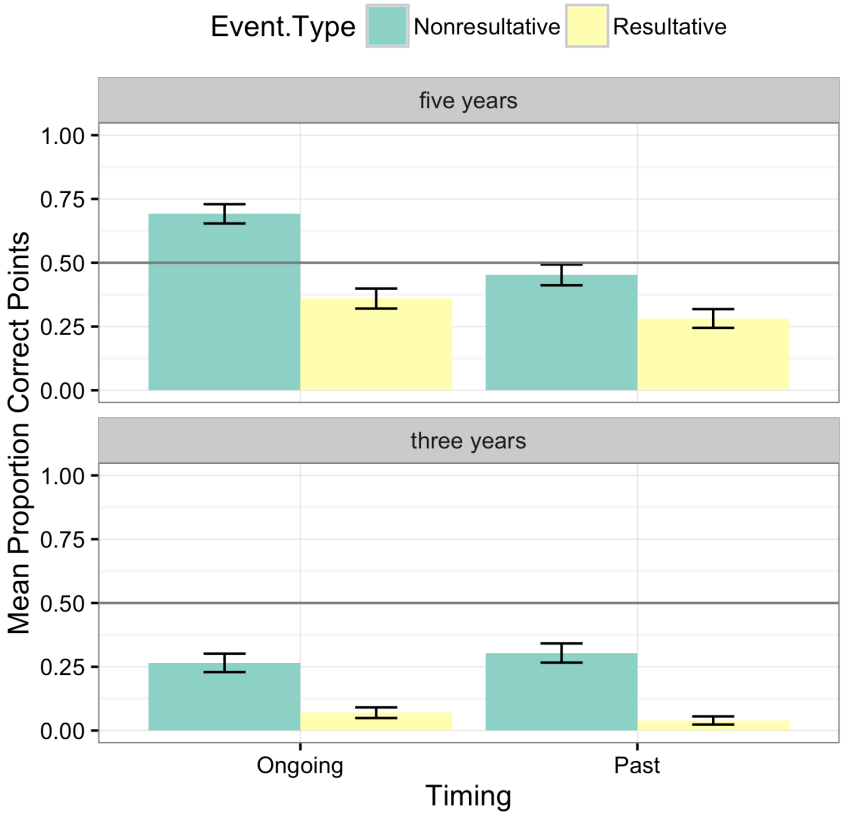
Results and Discussion

One percent of trials were excluded from analyses because the child pointed to both video clips. As for Study 1, to compare performance across our two experimental conditions, we again used a binomial mixed effect models with effect coded factors (e.g. Baayan, Davidson, & Bates, 2008), whereby the factors were Event Type (Resultative vs. Nonresultative), Timing (Ongoing vs. Past) and Age (3-year-olds vs. 5-year-olds). Participants were treated as random effects, with random slopes for participants. The *p*-

values were computed by comparing models with likelihood-ratio tests and chi-square values are reported. There was a main effect of Age Group (b = -4.09, SE = 0.88, $\chi^2(1)$ = 24.86, $p < .001$, showing that the five-year-olds overall were more likely to point at the target (Same Action) clip than were the three-year-olds. There was a main effect of Event Type (b = -3.87, SE = 1.36 $\chi^2(1) = 9.47$, $p < .01$), reflecting worse performance on the Resultative conditions than the Non-Resultative conditions. All other main effects and interactions were not significant.

Since we were specifically interested in following up previously established developmental effects, we also carried out binomial mixed effects models for each Age Group separately, with Event Type and Timing fully crossed. For both Age Groups, the only significant effect was for Event Type (b = -6.07, SE = 2.70, $\chi^2(1) = 9.33$, $p < .01$ for three year olds; b = -3.44, SE = 1.40, $\chi^2(1) = 7.65$, $p < .01$ for five year olds), whereby both age groups performed worse with the Resultative events. Neither age group showed a significant interaction between Event Type and Timing, nor a main effect for Timing (although the latter showed a trend in the direction of significance for the five-year-olds, b = -1.69, SE = 1.03, $\chi^2(1) = 2.90$, $p = .09$).

Fig 2 Mean proportion correct points for Study 2, by Event Type, Age Group and Timing Condition.



We also investigated whether the children performed above chance in any of the conditions, whereby the dependent variable was the proportion of target points, conflated over the six experimental item trials. Figure 2 above shows the mean proportion of correct responses (i.e. points to the Target clip), by Age Group and condition, with the grey line indicating chance level performance. The only condition in which any age group performed above chance was that of the five-year-olds (M = 69%

correct) with the non-resultative events when the novel verb was heard whilst the action was ongoing (in the present progressive tense) (t (19) = 2.44, $p < .05$). Since the original study (Imai et al., 2008) also used present progressive paired with ongoing actions and used non-resultative events, Study 2 (like Study 1) replicated the original results for English-speaking five-year-olds. The three-year-olds pointed significantly below chance in all conditions (all $p < .05$) suggesting they interpreted the novel word as relating to the object rather than the action. The five-year-olds only pointed significantly below chance in the condition where they saw resultative events and were taught and tested on the novel verb after the action was completed (t (22) = 2.73, $p < .05$), which ran precisely counter to our prediction that this would be the condition in which children performed best. Verb meanings were not learned more easily when they were taught on completion of the action, for any Age Group or Event Type. Three-year-olds and even five-year-olds in certain contexts appear willing to map a novel verb heard in an active transitive (e.g. 'She's mooping it') onto a novel object.

General discussion

We carried out two studies to replicate and extend Imai et al.'s (2005, 2008) paradigm, which pits novel objects versus novel actions as potential referents for novel verb learning. In line with Imai et al. (2005, 2008) we found that five-year-olds were able to choose the correct clip (i.e. blicking = action) for the Non-Resultative Event Type (e.g.

iterative 'rolling-between-palms' or 'fencing-stabs', see Appendix 1). Over both studies, children found the resultative events (e.g. 'flipping over' or 'head-butting-away', see Appendix 2) more difficult than the original non-resultative events. In Study 2 we found that hearing the new verb in the past tense (e.g. *blicked*) after the action was complete did not have the predicted ameliorating effect on how successfully either five- or three-year-olds mapped the new verb for resultative events.

Our findings from Study 2 do not fit with previous findings (e.g. Tomasello & Kruger, 1992; Ambalu et al., 1997; Carr & Johnston, 2001) which had suggested that teaching verbs on completion of the action would lead to better performance with resultative event types than teaching verbs whilst the action is still ongoing. To the contrary, both age groups performed poorly in our past conditions; the difference between the past and ongoing conditions was of marginal significance for the five-year-olds. However, it is possible that if we had used change-of-state actions, we would have found the predicted interaction between tense and event type. That said, there is no evidence that change-of-state events are more prototypically causative than the change-of-location events that we used; the evidence that exists, although sparse, appears to suggest that both are in fact prototypically causative as long as the action is intentionally caused (e.g. Muentener & Lakusta, 2011). Moreover, there are a limited number of *novel* ways in which a state can change (i.e. so that a pre-schooler would not simply describe the event as *break*, *clean*, *colour*, *cover*, *fix* or *open*/*shut*) and once a

novel object has changed state, it is then not the same object as it was initially, which might be problematic for the Same-Object foil used in the original studies (e.g. Imai et al., 2005; 2008). Furthermore, Tomasello and Kruger (1992) used a resultative action involving a punctual, change of location and this was learnt better when taught on completion on the action than when taught whilst the action was still ongoing.

While we clearly replicated Imai et al.'s (2008) findings for English-speaking five-year-olds, this is less clear for the three-year-old groups, who performed significantly below chance even in the replication condition (resultatives with ongoing action), indicating that they were mapping the novel verb in '*She is <u>blicking</u> it*' onto the novel object. The English-speaking three-year-olds in Imai et al.'s (2008) study pointed to the target (Same-Action) clip 42% of the time, which was not significantly below chance. That said, with the same stimuli Imai et al (2008) found that three-year-olds had very low performance; Japanese-speaking three-year-olds were correct 39% of the time in the Verb condition. For Mandarin-speaking children, even five-year-olds were only correct 17% of the time in the Verb condition, which was significantly below chance.

Therefore, our findings clearly line up with those of Imai et al. (2008) and also Kerstin and Smith (2002) to indicate that three-year-olds will, when given a choice between a novel action and a novel object, frequently assume that the novel word refers to the novel object. The forced choice between a novel object and a novel action is particularly difficult because it is a true 'Quinean' (Quine, 1960) scenario and because

objects appear to have much greater salience than the Same Action (e.g. Gentner, 1982, 2006; Kim et al., 2000; Markman, 1991), presumably due to factors such as temporal permanence, greater concreteness, individuation and imageability (e.g. Golinkoff & Hirsh-Pasek, 2008; McDonough, Song, Hirsh-Pasek, Golinkoff, & Lannon, 2011).

We cannot, however, determine exactly why the presence of a novel object makes the process of fast-mapping novel verbs to novel actions more difficult for young children. While it could be that children have a bias to map novel words onto whole objects (e.g. Markman, 1991), it is also possible that the pattern of results could be at least partially due to the difficulties which three-year-olds face with response inhibition (e.g. Beck, Schaefer, Pang & Carlson, 2011). That is, the greater concreteness, individuation and temporal permanence of objects may attract children's attention, lead to a greater depth of encoding in memory and thus when presented with a forced choice, they find it difficult to inhibit the inclination to select the novel object. Future studies could potentially attempt to disentangle these possibilities by investigating whether performance in this paradigm correlates with measures of inhibitory control (see e.g. Krott & Snape, 2015, for this type of relationship for a different word learning paradigm). Whatever the underlying reason, the predominant factor may be that of relative visual salience of particular items. We used change of location actions which are frequently punctual events. These might in fact be particularly difficult to fast-map new verbs onto simply because their temporal brevity makes their details (at least of

manner of motion) more difficult to encode in memory than durative actions. Interestingly, punctual change-of-location verbs such as *drop*, *throw*, *dump*, *knock over* are amongst the first handful of verbs which young English-speaking children learn (e.g. Fenson et al., 1994; Just, Christopher, Meints, Rowland, & Alcock, 2015). However, the addition of discourse and other socio-cognitive cues, such as cues allowing intention-reading, appears to allow even very young two-year-olds to map novel punctual, change of location causative actions onto novel words in the presence of novel objects (Tomasello & Akthar, 1995; see also Hohenstein, 2013, for a constructivist account of how parent-child interaction scaffolds the acquisition of motion verb semantics).

In sum, it appears that basic perceptual components of actions play a predominant role in the fast-mapping phase of the acquisition of verb meaning. While a role for perceptual salience is built in certain theories of word learning such as the Emergentist Coalition Model (e.g. Golinkoff & Hirsh-Pasek, 2008), what this theory does not predict is that perceptual salience continues to play such a predominant role at three and five years that it overrides morpho-syntactic cues. (Notably in the current study, the novel verbs were heard in an active transitive frame 'SUBJECT is VERBing OBJECT' with a case-marked subject). Therefore, in the absence of socio-cognitive cues (e.g. Tomasello & Akthar, 1995) objects do appear to be more visually salient than actions and this perceptual salience appears to swamp the syntactic knowledge of three-

year-olds when visual perceptual salience and morpho-syntax are pitted against one another.

Conclusion

Our findings fit with a large body of evidence that young children find it easier to learn new words for objects than for actions (e.g. Childers & Tomasello, 2002, 2006; Gentner, 2006). In the absence of social-cognitive cues, three- and sometimes even five-year-olds appear influenced by visual perceptual salience factors to a greater degree than by prototypical causality when fast-mapping novel verbs. This may lead them to frequently (initially) mistakenly fast-map novel verbs onto novel objects, because objects are more salient than actions, and it also may lead them to have more difficulty mapping those actions which are of temporal brevity. Since children must learn a large number of punctual verbs, our results elucidate something of the nature of the complexity of the task which children face when they hear new verbs for the very first time.

Acknowledgments
A very special thanks to Amanda Norman for testing children and for organising appointments. Thank you also to Natalie Fairhurst, Natalie Rusher and Precious Sango for collecting some of the data; to Nuffield foundation SGS/37511 to the first author for supporting the research and to Anthony Mee (University of Plymouth) and Jason Marshall for assistance with programming. Samantha Durrant is a Research Associate in the International Centre for Language and Communicative Development (LuCiD) at the University of Liverpool. The support of the Economic and Social Research Council [ES/L008955/1] is gratefully acknowledged. A special 'thank you' to the parents who brought their children into the Kent Child Development Unit and to the staff and caregivers of the following schools: Ashford Oaks Primary School, Ashwell Primary, Blean Preschool, Bournehall Primary, Chilham St. Mary's Preschool, East Stour Primary School, Highview Primary, Langland Community School, Lilliput Preschool, Little Drapers Nursery, Mapel Grove Primary, Petham Primary School, St. James the Great Primary, Standens Barn Preschool, Weston Favell Preschool and Wye Preschool. Many thanks also to Franklin Chang (LuCiD) for an excellent course on Linear Models in R. Please contact the first author on K.Abbot-Smith@kent.ac.uk to access the anonymised data files and scripts.

References

Ambalu, D., Chiat, S., & Pring, T. (1997). When is it best to hear a verb? The effects of the timing and focus of verb models on children's learning of verbs. *Journal of Child Language, 24*(1), 25-34. doi:10.1017/S0305000996002978

Arunachalam, S. & Waxman, S. (2011). Grammatical form and semantic context in verb learning. *Language Learning and Development, 7*, 169-184. doi: 10.1080/15475441.2011.573760

Arunachalam, S. & Waxman, S. (2015). Let's see a boy and a balloon: argument labels and syntactic frame in verb learning. *Language Acquisition, 22*(2), 117-131. doi: 10.1080/10489223.2014.928300

Baayen, R.H., Davidson, D.J. & Bates, D.M. (2008). Mixed-effects modelling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390-412. doi:10.1016/j.jml.2007.12.005

Bahrick, LE, Gogate, LJ & Ruiz I. (2002). Attention and memory for faces and actions in infancy: The salience of actions over faces in dynamic events. *Child Development, 73*, 1629-1643. doi: 10.1111/1467-8624.00495

Beck, D., Schaefer, C., Pang, K. & Carlson, S. (2011). Executive function in preschool children: test-retest reliability. *Journal of Cognition and Development, 12*(2), 169-193. doi:10.1080/15248372.2011.563485

Brandone, C. A., Pence, K. L., Golinkoff, R. M., & Hirsh-Pasek, K. (2007). Action speaks louder than words: young children differentially weight perceptual, social, and linguistic cues to learn verbs. *Child Development, 78*(4), 1322-1342. doi:10.1111/j.1467-8624.2007.01068.x

Carr, L., & Johnston, J. (2001). Morphological cues to verb meaning. *Applied Psycholinguistics*, *22*(4), 601-618. doi: 10.1017/S0142716401004064

Chan, C., Tardif, T., Chen, J., Pulverman, R., Zhu, L., & Meng, X. (2011). English- and Chinese-learning infants map novel labels to objects and actions differently. *Developmental Psychology*, 47(5), 1459-1471. doi: 10.1037/a0024049

Childers, J. B., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology, 38,* 967–978. doi:10.1037//0012-1649.38.6.967

Childers, J. & Tomasello, M. (2006). Are nouns easier to learn than verbs? Three experimental studies. In K. Hirsh-Pasek & R. Golinkoff (Eds.). *Action meets word: how children learn verbs*. Oxford University Press.

Fenson, L., Dale, P., Reznick, J., Bates, E., Thal, D. & Pethick, S. (1994). Variability in early communicative development. *Monographs for the society for research in child development*, 59 (5). Chicago, IL: University of Chicago Press. doi:10.2307/1166093

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. A. Kuczaj (Ed.), *Language development: Vol. 2. Language,*

thought, and culture (pp. 301 – 334). Hillsdale, NJ: Erlbaum.

Gentner, D. (2006). Why verbs are hard to learn. In K. Hirsh-Pasek & R. Golinkoff (Eds.). *Action meets word: how children learn verbs*. Oxford University Press.

Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning words and rules: Abstract knowledge of word order in early sentence comprehension. *Psychological Science, 17*(8), 684 - 691. doi: 10.1111/j.1467-9280.2006.01767.x

Golinkoff, R., & Hirsh-Pasek, K. (2008). How toddlers begin to learn verbs. *Trends in Cognitive Sciences, 12*(10), 397-403. doi: 10.1016/j.tics.2008.07.003

Gropen, J., Pinker, S., Hollander, M., & Goldberg, R. (1991). Affectedness and direct objects: the role of lexical semantics in the acquisition of verb argument structure. *Cognition, 41*, 153-195. doi:10.1016/0010-0277(91)90035-3

Hohenstein, J. (2013) Parent–child talk about motion: Links to children's development of motion event language. *First Language 33*, 411-425. doi 10.1177/0142723713490599

Imai, M., Haryu, E., & Okada, H. (2005). Mapping novel nouns and verbs onto dynamic action events: Are verb meanings easier to learn than noun meanings for Japanese children? *Child Development*, *76*(2), 340-355. doi: 10.1111/j.1467-8624.2005.00849_a.x

Imai, M., Li, L., Haryu, E., Okada, H., Hirsh-Pasek, K., Golinkoff, R. M., & Shigematsu, J. (2008). Novel Noun and Verb Learning in Chinese-, English-, and

Japanese-Speaking Children. *Child Development*, *79*(4), 979-1000. doi: 10.1111/j.1467-8624.2008.01171.x

Just, J., Christopher, A., Meints, K., Rowland, C. F., & Alcock, K. J. (2015). The effect of age on the composition of the first 10 words: Evidence from the UK-CDI. *Child Language Symposium, University of Warwick, UK, July 2015.*

Kersten, A. & Smith, L. (2002). Attention to novel objects during verb learning. Child *Development, 73*(1), 93-109. doi: 10.1111/1467-8624.00394

Kim, M., McGregor, K., & Thompson, C. (2000). Early lexical development in English- and Korean-speaking children: language-general and language-specific patterns. *Journal of Child Language, 27*, 225-254. doi: 10.1017/S0305000900004104

Krott, A. & Snape, S. (2015). The role of inhibition in structural alignment. *Child Language Symposium, University of Warwick, 20th -21st July 2015.*

Liss, P. & Haith, M. (1970). The speed of visual processing in children and adults: effects of backward and forward masking. *Perception & Psychophysics, 8*, 396-398. doi: 10.3758/BF03207032

Maguire, M., Hirsh-Pasek, K., Golinkoff, R., & Brandone, A. (2008). Focusing on the relation: fewer exemplars facilitate children's initial verb learning and extension. *Developmental Science, 11*(4), 628-634. doi:10.1111/j.1467-7687.2008.00707.x

Markman, E. (1994). Constraints on word meaning in early language acquisition. *Lingua, 92*, 199-227. doi:10.1016/0024-3841(94)90342-5

Markman, E. M. (1991). The whole-object, taxonomic, and mutual exclusivity assumptions as initial constraints on word meanings. In S. A. Gelman, J. P. Byrnes (Eds.), *Perspectives on language and thought: Interrelations in development* (pp. 72-106). New York, NY US: Cambridge University Press. doi:10.1017/CBO9780511983689.004

Markson, L., & Bloom, P. (1997). Evidence against a dedicated system for word learning in children. *Nature*, 385, 813–815. doi: 10.1038/385813a0

McDonough, D., Song, L., Hirsh-Pasek, K., Golinkoff, R. & Lannon, R. (2011). An image is worth a thousand words: why nouns tend to dominate verbs in early word learning. *Developmental Science, 14*(2), 181-189. doi:10.1111/j.1467-7687.2010.00968.x.

Meints, K. (1999). *Typikalizitätseffekte im Erwerb des englischen Passiv.* Opladen: Westdeutscher Verlag.

Muentener, P. & Lakusta, L. (2011). The intention-to-CAUSE bias: evidence from children's causal language. *Cognition, 119*(3), 341-355. doi:10.1016/j.cognition.2011.01.017

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language* 17, 357-374 doi: 10.1017/S0305000900013817

Noble, C. H., Rowland, C. F., & Pine, J. M. (2011). Comprehension of argument structure and semantic roles: Evidence from English-learning children and the

forced-choice pointing paradigm. *Cognitive Science, 35*(5), 963-982. doi: 10.1111/j.1551-6709.2011.01175.x

Oshima-Takane, Y., Ariyama, J., Kobayashi, T., Katerelos, M. & Poulin-Dubois, D. (2011). Early verb learning in 20-month-old Japanese-speaking children. *Journal of Child Language*, 38, 455-484. doi: 10.1017/S0305000910000127

Quine, W. V. (1960). *Word and object.* Cambridge, MA : MIT Press.

Scott, R. & Fisher, C. (2009). Two-year-olds use distributional cues to interpret transitivity-alternating verbs. *Language and Cognitive Processes*, *24*(6): 777-803 doi: 10.1080/01690960802573236

Semel, E., Wiig, E. H. & Secord, W. (2004). *Clinical Evaluation of Language Fundamentals- Preschool-2* (2nd ed.). San Antonio, TX: NCS Pearson Inc.

Slobin, D. (1981). The origins of grammatical encoding of events. In W. Deutsch (Ed.). *The child's construction of language*. P. 185-199. London: Academic Press.

Tardif, T. (1996). Nouns are not always learned before verbs: evidence from Mandarin speakers' early vocabularies. *Developmental Psychology, 32*(3), 492-504. doi: 10.1037/0012-1649.32.3.492

Tomasello, M., & Kruger, A. C. (1992). Joint attention on actions: Acquiring verbs in ostensive and non-ostensive contexts. *Journal of Child Language*, *19*(2), 311-333. doi: 10.1017/S0305000900011430

Tomasello, M. & Akhtar, N. (1995). Two-year-olds use pragmatic cues to differentiate reference to objects and actions. *Cognitive Development, 10*(2), 201-224. doi:10.1016/0885-2014(95)90009-8

Appendix 1: Description of actions and objects for Non-resultative (replication) condition

Target action	Same Action	Same Object
Palm-roll	Action = *palm roll* (Rolls object between palms of hands) Object = has long handle and bulb-like piece on end	Action = *repeatedly lightly toss and catch* Object = American-football-shaped object with coloured stripes and purple 'feet' parts
Dry-back	Action = *Back-dry* (Hold object behind back & pull it up and down as if drying your back) Object = Long, narrow meshed grey twisty object.	Action: *Torso-twist* (Holds object in front of self, with a hand at each end, and twist torso from side to side). Object: A long, blue, plastic arch-shaped ball-thrower for dogs.
Punch-outwards	Action*: crucifix-defence* (holds object in right hand and push it outward as if holding a crucifix to ward off a vampire). Object: A round metallic	Action: *Shoulder-tap* (Holds object in right hand and taps against left shoulder). Object: A black plastic angular drainpipe part with red stripes.

	timer.	
Wring-out	Action: *Wring out* (Hold object in both hands at chest level, twisting it so that it bends in the middle, as if wringing out a wet cloth); Object: A large circular blue rubber ring.	Action: *Flop up and down* (Holds object in right hand and slightly move that hand so that the object flops up and down). Object: A large red rubber dog toy consisting of two loops.
Fencing-pronation-tierce	Action: *Repeated fencing-stabs* (hold object in right hand, pushing it out with a stabbing motion). Object: A long wire cylindrical CD rack	Action: *Knee-tap* (Holds object in right hand and tap against her right knee, which she raises at the same time as she is lowering the object). Object: A bamboo candle holder on a long thin stem.
Thigh-tap	Action: *Thigh-tap* (tap against thigh) Object: A large black curved pipe	Action: *Finger-roll* (Hold object with the index fingers of each hand stuck into the hole in the centre and roll the object around fingers). Object: A round wooden toy with holes in the top for shapes to be put into.

Appendix 2: Description of actions and objects for Resultative condition

Target action	Same Action	Same Object
Flip-over	Action: *Flip over* (put hand on edge of object and flips it over so that it lands upside-down on the other side of the table) Object: Large, blue, round, plastic object with legs.	Action: *Finger-twirl-and-toss* (twirls object around index finger so that it flies off). Object: Small, round object made of metal with criss-crossing parts across the radius and a green plastic rim with green fins.
Foot-drop	Action: *Foot-drop* (balance object on foot and then withdraws foot so that object drops). Object: A blue plastic bulb with a yellow base and round white plate on top.	Action: *Bullwinkle-antler-set* (balance object on the fingertips of both hands and then tosses it up into the air like a Bullwinkle-antler set in volleyball). Object: A large light blue plastic square object.
Dog-throw	Action: *Dog-mouth-throw* (Hold object in mouth and then throw by tossing head).	Action: *Elbow-jerk* (balance on elbow and then toss upwards by jerking elbow).

	Object: A large round wicker clothes basket lid.	Object: A red round plastic half of a swing-top dustbin lid.
Underhand-serve	Action: *volleyball-underhand-serve* (balance object on fist and bring other fist underneath the first in a punching motion). Object: solid metal cylindrical object decorated with red flowery bows	Action: *Finger-flick* (prototypical interpretation of flick) Object: A wooden rectangular object with three round holes in the side and a red handle.
Hammer-throw	Action: *Olympic-hammer-throw* (swirl once around head and then throw like an Olympic hammer-thrower). Object: Metal toilet roll stand with spiral-shaped metal base and covered in red ribbons.	Action: *Upwards-wrist-flick* (Hold object between index fingers and thumbs by both hands and flick it up into the air). Object: Blue plastic oblong object with two cross-pieces.
Head-butt	Action: *head-butt* (Holds object in both hands and head-butt it onto the table).	Action: *Elbow* (Elbow object so that it topples over). Object: A stacking ring with white

	Object: A red, yellow, blue and green striped kite rolled up so that it is long and thin but with streamers hanging off it.	plastic base and yellow plastic stem.