Long-tailed macaques select mass of stone tools according to food type

Michael D. Gumert1 and Suchinda Malaivijitnond2

1Division of Psychology, School of Humanities and Social Sciences, Nanyang Technological University, Singapore 637332, Republic of Singapore
2Primate Research Unit, Department of Biology, Faculty of Science, Chulalongkorn University, Bangkok, Thailand

Tool selection can affect the success of a tool-based feeding task, and thus tool-using animals should select appropriate tools when processing foods. We performed a field experiment on Piak Nam Yai Island in Laem Son National Park, Thailand, to test whether Burmese long-tailed macaques (Macaca fascicularis aurea) selected stone tools according to food type. We baited the island’s shores with stone sets (‘tool tests’) in an effort to attract macaques to use stones presented in a quasi-experimental design. Tool tests were placed at 344 locations for 126 days over a 2 year period, with each set containing four stones of different mass (categories: X, 40–60 g; S, 90–100 g; M, 150–200 g; and L, 400–1000 g). Tool tests were checked when we could access them. The number of times each tool test was checked varied (1–32), for a total of 1950 checks. We also studied 375 non-experimental stone tools that were found at naturally occurring tool-use sites. Our data were not collected by direct observation, but by inspecting stones after use. We found an association between stone mass and food type. In the tool tests, we found S-stones were chosen most often for attached oysters, and L-stones were chosen most often for unattached foods. L-stones were almost always chosen for larger unattached foods (greater than 3 cm length), while for smaller unattached foods (less than or equal to 3 cm length) selection was less skewed to L-stones and more evenly distributed between the M- and L-stone categories. In the non-experimental study, we found that mass varied significantly across five food categories (range: 16–5166 g). We reveal more detail on macaque stone tool mass than previous studies, showing that macaques select differing stone masses across a variety of tool-processed foods. Our study is the first step in investigating the behavioural and cognitive mechanisms that macaques are using during tool selection.

1. Introduction

Animals that use tools would be aided by an ability to discriminate material that would make the most suitable tool for the task. Having such selective capacities would reduce the costs associated with superfluously using non-functional tools [1–3]. An ability to recognize the use of an object, or its affordances [4], is a central feature underlying the cognition driving human technology. For example, archaeologists show that prehistoric humans selected appropriate stones for knapping and making stone flakes [5]. Stone pounding is a probable precursor of making stone tools [6,7], and therefore understanding stone tool use in extant stone-using non-human primates provides a foundation for investigating the evolution of human tool manufacture and use [3].

Several animal species select objects as tools on the basis of specific features. For example, New Caledonian crows (Corvus moneduloides) and woodpecker finches (Camarhynchus pallidus) select sticks of appropriate shape, length and diameter to insert into holes to obtain food [8–14]. Egyptian vultures (Neophron percnopterus) [15] and black-breasted buzzards (Hamirostra melanosternon) [16] select stones used to drop on eggs within a certain range of mass. Captive studies show that chimpanzees (Pan troglodytes) [17], capuchins (Cebus apella) [18], orangutans (Pongo pygmaeus) and gorillas (Gorilla gorilla) [19] select the most functional sticks to retrieve food. Similarly, wild chimpanzees select...
functional stick tools, by choosing sticks of appropriate size and length for honey dipping and termite fishing [20–22]. Wild chimpanzees select heavier and harder stones to process more resistant nuts [23], whereas both captive chimpanzees and capuchins select hammers based on mass [24,25]. Visalberghi et al. [3] and Fragaszy et al. [1] have studied stone selection in wild bearded capuchins (Sapajus libidinosus) using field experiments, and demonstrated their ability to select the most functional material and mass for opening nuts [26]. Systematic field observations have complemented these findings, confirming that bearded capuchins naturally select for stone properties according to the resistance of the encased foods processed [27,28].

Burmese long-tailed macaques (Macaca fascicularis aurora) use stone tools along the western coast of Thailand [29] and Myanmar [30]. At Laem Son National Park, they use two types of stone tools: small hand-sized stones to process sessile oysters; and larger, heavier, harder stones to process collected nuts and shellfish on anvils [31]. A recent study investigating the foods processed by long-tailed macaques found they processed over 30 different mollusc species, several species of crustaceans and other animals, and the nuts of four coastal plant species [32]. With this new information on food species that macaques process with tools, it is now possible to study how stone features vary across them, rather than just comparing between the classes of sessile or unattached foods.

We conducted a study on Piak Nam Yai Island in Laem Son National Park in Thailand to investigate how stone mass is related to the various encased foods that macaques process with stones. We conducted a field experiment and also studied stones found at natural tool-use sites. For the field experiment, we placed sets of stones, or ‘tool tests’, on the shores of Piak Nam Yai. Each tool test contained four stones of different mass. We then investigated how macaques selected and used the four stones for processing different species of food. During the field observations, we visited areas of coast where we had observed macaque tool use. During these visits, we identified each food species from the remaining debris and measured the mass of stone tools found at each tool-use site.

The types of food that macaques process with stones vary in their features [32], so we hypothesized that macaques would actively select for stone mass according to the food being processed. Based on previous findings [31], we predicted that stone mass would vary with the mobility of the food, and that smaller, lighter stones would be more strongly associated with sessile oysters, whereas larger, heavier stones would be associated with motile and detached foods. In addition, given the wide variety of motile prey, as well as some nuts, being processed by macaques [32], we were able to test how stone tool mass relates to foods in greater detail than in past studies. We expected that stone mass would not only vary based on the mobility, but also with the size and toughness of the food. We predicted among motile prey and nuts, that larger, and more resistant foods would be processed with stones of larger mass, whereas smaller, less resistant foods should be processed with stones of smaller mass.

2. Methods

(a) Field site

We conducted this study on the coast of Piak Nam Yai, which is a small island located approximately 500 m from the mainland in Laem Son National Park, Ranong Province, Thailand. Piak Nam Yai is one of 15 protected islands in Laem Son, is 1.7 km² in size, and has 3.4 km of coastline. The shoreline is composed of 66% rocky shore, 29% mangrove and 5% sandy beach. These intertidal regions are affected daily by tides that can vary as much as 4 m during spring tides, posing considerable challenges for fieldwork. We accessed the intertidal areas of Piak Nam Yai by boat. Gumert et al. [33] described the features of the park and research site in greater detail.

The macaques used stones as tools in the rocky shore and mangrove intertidal regions. Anvils were clearly distinguishable from tools, and we never observed them being moved by macaques. The rocky shores consisted of large boulders, spanning up to several metres across. The expansive flat surfaces of these boulders provided the most common anvil surfaces on which macaques pounded food items. In the mangroves, large boulders were much less common, and anvils were generally rocks spanning ca 0.5–1 m across, with a small proportion of rocks up to 3–5 m across. Stone tools were generally found on these large rock surfaces, or within oyster beds.

(b) Macaque population

Piak Nam Yai has ca 200 macaques living in nine groups, ranging from 8 to 35 individuals per group [33]. The macaques are continuously distributed along the coast of Piak Nam Yai and regularly forage in the intertidal zones where they are exposed. Gumert et al. [33] identified all adult and adolescents on Piak Nam Yai in 2011 (n = 107). All nine macaque groups come to the shores, and their tool use is highly conspicuous. These macaques are customary tool users, as defined by McGrew [34], Whiten et al. [35] and Shumaker et al. [36]. We have determined that 88% of all adolescents and adults living on Piak Nam Yai use stones daily to process foods [33], which amounts to ca 20% of their total activity when on the shore, and 40% of their coastal foraging time budget (M.D. Gumert 2009–2012, unpublished data).

The macaques were unhabituated and foraged in difficult to traverse coastal terrain, so they could only be directly observed at a distance (ca 10–50 m) from a boat (i.e. Thai long-tail boat, approx. 12 m long x 2.5 m wide). Through these observations, we have acquired a database of their basic activities and general patterns of tool use (M. D. Gumert 2007–2012, unpublished data). Although we can easily see macaques and can watch them for hours each day, we have difficulty in acquiring information on the details of their tools, foods and behaviour through observation alone, due to the distance, large boulders on the shores affecting visibility, and wave motion affecting the stability of the boat. We therefore have relied heavily on indirect techniques to obtain more detailed information of their activities. The data collected were not from behavioural observations of macaques. Rather, we inspected tool tests and natural tool-use sites after their use.

(c) Experimental design of tool tests

We presented tool tests on Piak Nam Yai’s shores between December 2009 and June 2011 in three phases for a total of 126 days. Phase 1 occurred between 7 and 21 December 2009; phase 2 between 21 February and 10 March 2011; and phase 3 between 23 March and 23 June 2011. We set tool tests at 344 different locations (figure 1). During phase 1, we explored how to present tests for use by macaques, and therefore, set many more tests than in the latter two phases.

Rocky shore and mangroves tool tests differed in their placement in the intertidal zone. On the rocky shores, tool sets were set above the high tide mark, and thus pounded food debris was easy to recover. Stones that macaques moved lower into the littoral were easily washed away and lost, along with any food debris, and thus were difficult to recover. As a result, our ability to study oyster tool use from tool tests presented on rocky
shores was limited. In the mangroves, it was easier to test oyster tool use, because the tests could be presented in areas that were submersed during high tide, and oysters grow in the regions submersed by tides. However, being submersed in the tides made it harder to test stones used for pounding, because debris on anvils was easily washed away. We did not aim to investigate differences between the two environments, as there were obvious potential biases in each condition based on how we had to set the tests. Rather, each habitat type had advantages for testing certain food species (i.e. mangrove–oysters; rocky shore–sea snails and nuts). Therefore, by placing tools in both environments, we increased our ability to have tool tests used for a wider variety of food species and better test how stone mass related to food type.

We designed tool tests consisting of four stones that varied within the range of mass for tools most commonly used by macaques (ca 40–1000 g) [31]. Each test had one stone in each of the following categories: extra small (X; 40–60 g), small (S; 90–100 g), medium (M; 150–200 g) and large (L; 400–1000 g; figure 1). Because size (i.e. volume) also varies with the mass of objects (i.e. amount of matter) that are of similar material and density, these stones also varied in size. There was no way to separate these two physical characteristics of the stones, and thus this was not a study to discriminate whether macaques attended to either one of these characteristics more when selecting a stone tool. Based on data from past studies, mass, volume and several other physical dimensions of macaque tools all correlate [31], and thus any of these measures can serve as a general descriptor of the size of a macaque stone tool.

We collected stones on Piak Nam Yai’s shores and measured their mass using a portable electronic scale (VWR CLW5000). We then sorted stones into the four categories and discarded stones with an unfitting mass. We cleaned each stone to remove any dirt and debris, dried the stones in the sun, and then used paint markers (UniPaint PX-30) to mark them with a number (figure 1). Each category was marked with a different colour so they were easier for us to identify from a distance. X stones were marked with blue, S in black, M in pink and L in green. When setting tool tests, we matched stones together that were visibly similar in rock type, shape and surface texture to control for these characteristics. We also painted pink and blue patches on many stones to help us see and relocate moved stones among the many shore rocks.

We set tool tests at various locations in the mangroves and rocky shores. Tests were set mainly during low tide when the intertidal regions were exposed, and we could land on the shore and walk through. At times, we also snorkelled through the mangroves in high tide for test setting, and this was done to increase time efficiency and to avoid deterring the macaques from using the area or disturbing the mud flats.

Tool tests were placed in the areas based on the following three criteria: the ease with which we could access the area, if tool use had been previously observed in the area, and the suitability of the location to place stones without them being significantly disturbed by the sea. We chose 11 different coastal areas to place our tests (figure 1) and they varied between 50 and 250 m long. Four areas were in the mangroves and seven on the rocky shores, but total shore length tested was similar for the two habitats. After an area was selected, we placed five to 60 tool tests on flat stones, the ground or in baskets. We placed the four stones of each test in a straight line, in order of mass category, so that we could quickly and easily observe later if they had been displaced from their original position (figure 1a). The baskets set-up in the mangrove were attached to tree roots or rocks with cable ties and garden wiring (figure 1b), and in the baskets stones were not placed in a straight line.

Figure 1. A map of Piak Nam Yai Island showing the location of the 11 test areas (labelled) and 344 tool test locations (filled circles). Thin grey borders demark a sandy beach on the western coast, and mangrove and mudflats regions on the eastern coast. R, rocky shore test areas, and M, mangrove test areas. The right panel shows (a) a tool test placed on a flat boulder on a rocky shore, (b) a tool test placed in a basket in a mangrove, and (c) a used tool test on a rocky shore with Thais bitubercularis debris. (Online version in colour.)
(d) Monitoring and scoring of tool tests

Test locations were recorded by GPS, and we subsequently returned to each test at least once to monitor its use. The quality of weather conditions and tidal level determined when we could access the tool test sites, and thus tests were not checked at standard time intervals. Checks usually occurred every day or two. However, intervals between checks could extend for as long as two weeks depending on the conditions (e.g. poor weather, water and safety conditions). During a check, the researcher walked, or occasionally snorkelled, along the shoreline and looked at each tool test. Quite often, the researcher was limited in time because of the rising tide. The researcher, therefore, had to be quick and selective in assessing the test and taking measurements, before water levels covered the tests, or lowered the safety of returning to the boat.

When a disturbed test (figure 1c) was located, the site was photographed and the attending researcher inspected what activity had occurred. A judgement had to be made if macaques or natural forces had disturbed the stones. It was possible to identify a tool from the following three features: (i) use-wear on the stone, (ii) shell dust adhered to the stone, and (iii) food debris around the stone (i.e. shell fragments, opercula, pigments, etc.). Sometimes, the food debris may have already been washed away or left at another location other than the stone’s final resting spot. In such cases, we had to rely on use-wear to infer tool use, however, such data were not used for assessing the relationship between tool mass and food. Macaques also moved stones with no clear evidence of use which could be identified by: (i) one stone being moved, whereas the others were not, (ii) stones were moved in different directions, or (iii) stones were found resting where wave or tide action could not have carried them (e.g. atop a higher boulder). In the absence of evidence of macaque activity, stones were considered disturbed by natural forces. Stones were judged as not moved by macaques when all stones were moved in the same general direction, suggesting that tide or wave action pushed them off their resting spot. Sometimes stones with no evidence of use were displaced small distances (less than 10 cm) and these were recorded as inconclusively disturbed. There were no other species on the island that would move stones for use, but agents that could have potentially displaced stones by inadvertent action included dogs, birds and monitor lizards. We had no evidence that humans ever disturbed or used our tool tests.

At used tool tests, we recorded the type of food debris, the distance the stones were moved and the patterns of use-wear. Use-wear scars were found on the points (i.e. short edges and protuberances), the edges (i.e. long edges) or the faces (i.e. flat sides) of the tool. Food debris found by the tool was collected, identified and each individual food item counted, and we also photographed the site. M.D.G. used the photographs and collected debris to later reconfirm all field recordings. The distance a stone was moved from its point of origin was measured using a tape measure. After all data were recorded, the stones were collected, wiped clean of dust and superficial scratches and returned to the tool test location. The tool test location was swept of any debris, and the test was reset.

A total of approximately 2000 different stones were used in our field experiment. We replaced stones only when they could not be recovered during a check, or were broken. Otherwise, displaced stones at tool tests were reset after a check. This method resulted in the repeated use of tool tests, and thus introduced pseudo-replication issues into our statistical procedures. However, this method was chosen for practicality and safety reasons, owing to the difficulties of collecting, carrying and placing stones. We were already using a very large number of stones from the island, and changing tool tests after every use would have significantly increased the number of stones required. Furthermore, we were also limited in our ability to safely alight from the boat and carry heavy buckets of replacement stones through the intertidal zones while performing checks. Last, this study was performed while engaged in several other studies, limiting time investment into tool test preparation.

(e) Analysis of tool tests

The tool test data were analysed to test how selectivity and movement related to stone mass and food type. We assigned a mass index number to each stone category (\(X = 0, S = 1, M = 2\) and \(L = 3\)) and tested if the mass index of stones selected by macaques was related to the food being processed. We could not compare the mass index across all food species, because we did not have enough data for each food. We therefore compared the mass index across three food groups based on the following general characteristics of food remains found with the stone. Group 1 contained stones with remains from sessile oysters, group 2 contained stones with debris from small species of gastropods that averaged less than or equal to 3 cm in length, including nerites and toothed-lip snails. Group 3 contained stones with debris from larger species of molluscs or nuts that averaged greater than 3 cm in length, including conches, drills, bivalves (i.e. clams, mussels and detached oysters), sea almonds and pandanus keys. The species in group 3 had tougher shells than other groups.

We chose to use non-parametric statistics, because the dataset was not that large and did not meet the assumptions of parametric statistics. We used Kruskal–Wallis tests to compare across food groups, and then used Mann–Whitney tests with Bonferroni corrections for paired post hoc comparisons. For all tests, alpha was set at 0.05.

For our tool selection analyses, we had some issues with independence of the data. First, we could not determine which individual macaques chose the stones for use, and thus the same individual could have chosen several stones. This would have produced repeated data for the same macaque’s tool use preferences. Second, in some tests checks, we found more than one tool had been used and hence some selections occurred after a tool test had already been disturbed. Third, we did not have sufficient capabilities or resources to change a tool test after each disturbance of the test, or to move it to a new location. We therefore reset the same test at the same location after disturbance. The consequence of this was that several tool tests had repeated selections taken from them (range = 1–12 selections per tool test, with a median = 1 and mean = 1.99). Because our dataset was not large, we were faced with the problem that reducing the dataset to account for repeated uses of some tool tests reduced the power to detect small differences between the food groups.

We ran our analysis in the following way. First, we made no correction, and ran the entire dataset of all tools to increase the power, using each point as an independent choice. We then ran a second test of the data using only tool tests where one tool had been found selected at a test check, which removed the influence of prior disturbance on stone selection. Last, to remove the independence bias of some tool tests being used repeatedly after being reset, we ran an analysis on a data subset of tool tests that had only one tool selected from them during the entire study. There was no way to control for the independence problem of individual macaques potentially using more than one tool and thus repeating data from an individual macaque. However, by analysing tests only used once, we did eliminate the possibility of getting repeated data from the same animal at the same tool test, and lowered the overall likelihood of one animal using more than one tool in the dataset.

For our analysis of stone movement, we also ran the first set of analyses using the entire dataset to obtain the full power of the data. Each moved stone was measured for the distance moved. We performed three comparisons on the displacement of the stones; (i) between stones used as tools and those just moved, (ii) across the four tool sizes, and (iii) across the three food
categories. In an effort to control for repeated sampling of the same-sized stone in each tool test, we also ran each analysis using only the first movement of each stone size in each tool test.

(f) Stone mass at natural tool sites
We measured the mass of all tools found with food debris (i.e. on same anvil, under or adjacent to tool or adhered to tool) during a concurrent study of food remains [32]. We measured all stones in an area immediately after observing a group of macaques using tools there. We collected 375 different stone tools, each at different locations, and categorized these stones by the type of food they had been used to process. This dataset was not normally distributed, but we did not have the independence issues with the stones in this analysis, other than we had no way of discerning which individual macaques used the stones. There was no way to remove this bias from the dataset. We used a Kruskal–Wallis test followed by Mann–Whitney post hoc paired comparison with Bonferroni corrections to test for difference in the tools across the food categories.

3. Results
(a) Summary of activity at tool tests
The total sum of test checks during the study was 1950 (mean = 6 checks per test, range = 1–32 checks). Phase 1 had 15 days of testing, 226 test locations and 1061 checks; phase 2, 18 days, 278 tests and 379 checks; phase 3, 93 days, 40 tests and 510 checks. In these 1950 test checks, we found 247 checks having evidence of macaque activity, of which 145 we inferred as tool use (figure 2a). From these 1950 tool test checks, there were 7800 individual stone checks (i.e. 1950 x 4). In these, we found 372 stones that showed evidence of macaque activity, of which 183 we inferred as being used as tools (figure 2b).

(b) Tool selection and food type
For 174 of the 183 tool use cases inferred during checks, we found recoverable food debris. The macaques processed 102 rock oysters (Saccostrea cucullata), 82 nerite snails (Nerita spp.), 42 drills (Thais spp. and Merula spp.), 14 toothed-lip snails (Monodonta labio), two venus clams (Gafarium spp.), one mussel (Perna viridis), one crown conch (Pugilina cochlidiun), two crabs (unidentifiable), one gecko lizard (Hemidactylus sp.), 11 sea almonds (Termiudia catappa), three rhizophora seedlings (Rhizophora apiculata), and two pandanus keys (Pandanus tectorius) during the experiment. More than one food item could occur at a single tool, and thus the number of individual food items is greater than 174.

We classified the various species of food into three groups. Group 1 (n = 66) was attached oysters, group 2 (n = 31) was small gastropods and group 3 (n = 57) was larger molluscs and nuts. There were 15 stones found with debris from multiple species (i.e. versatile tools), which were excluded because we could not assign a food group. Additionally, three stones were found with rhizophora seeds, one with crab debris and one with a crushed gecko. These rare cases were excluded, because we could not justify lumping them together or including them in one of the other food categories.

The three food groups significantly differed in their average food mass index score (Kruskal–Wallis: H_{2,58c} = 77.02, p < 0.001). Mann–Whitney post hoc tests indicated that in pairwise comparisons each group differed significantly from the other two (group 1: median = 1, mean = 0.97, s.d. = 0.78; group 2: median = 3, mean = 2.26, s.d. = 0.93; group 3: median = 3, mean = 2.65, s.d. = 0.77; p < 0.017). The distribution of tools selected from each mass category varied across food groups (figure 3). In particular, the selection pattern for group 3 was much more highly skewed towards L-stones than group 2, showing variation in selection patterns between two different food categories of unattached food. In group 3, 78.9% of all stones selected were L-stones, whereas, in group 3, L-stones only constituted 51.6% of the stones selected.

We analysed all tests checks that had only one stone selected, which allowed us to control for the effect of prior disturbance on selection. This gave us a sample where we were certain the macaque selected the one stone out of all four choices. We found that in 90 of the 145 used tests, there was only one tool used (62.1%). We analysed 76 of these tests that had recoverable food debris that could be categorized into the
three food groups. We found a significant difference in mass index across the three food groups (Kruskal–Wallis: $H_{2,76} = 37.13, p < 0.001$). After Bonferroni corrections ($\alpha = 0.017$), pairwise tests showed group 1 differed significantly from groups 2 and 3 ($p < 0.001$); however, for groups 2–3, the Bonferroni criteria were not met ($p = 0.046$; group 1: median = 1, mean = 0.92, s.d. = 0.76; group 2: median = 3, mean = 2.32, s.d. = 0.95; group 3: median = 3, mean = 2.69, s.d. = 0.76).

We conducted a final analysis to control for repeated use of the same tool tests with a subset of 48 stones from tool tests used only once in the study. Forty-two of these tests had a tool with food debris we could identify and categorize into one of the three food groups. After this test, we still found a significant difference in the mass index across the three food groups (Kruskal–Wallis: $H_{2,42} = 22.34, p < 0.001$), and post hoc tests showed group 1 significantly differed from groups 2 and 3 ($p < 0.001$). Groups 2 and 3 were not found to be different ($p = 0.869$). All three analyses produced the same pattern of results, but the winnowing of the data into smaller subsets removed the power to show a difference between groups 2 and 3.

(c) Stone movement

In addition to the 183 stones used, there were also 102 stones moved with no evidence of usage (i.e. no obvious scars and no food debris nearby; figure 2b). We analysed the movement patterns of all these 285 stone movements. We found that used stones were moved farther on average (median = 0.60 m, mean = 4.76 m, s.d. = 12.82 m) than stones that were only moved and not used (median = 0.30 m, mean = 2.02 m, s.d. = 5.48 m; Mann–Whitney: $n_1 = 183, n_2 = 102, U = 7149.50, p = 0.001$). Furthermore, we found that used stones had a larger average mass index (median = 2, mean = 1.93, s.d. = 1.11) than tools that were only moved, but not used (median = 1, mean = 1.30, s.d. = 1.04; $n_1 = 183, n_2 = 102, U = 6401.50, p < 0.001$). When we tested the dataset using the first movement of each stone size in each test, we found the same pattern of results, and thus we do not report them.

We tested whether the mass index affected the distance of movement and found a significant difference across the four mass categories (Kruskal–Wallis: $H_{3,285} = 31.72, p < 0.001$). Small (S) tools were moved the farthest (median = 0.90 m, mean = 8.11 m, s.d. = 17.77 m). Post hoc tests with Bonferroni corrections ($\alpha = 0.008$) showed S-tools were moved significantly farther than medium (M) tools ($p = 0.007$; median = 0.46 m, mean = 1.98 m, s.d. = 5.14 m) and large (L) tools ($p < 0.001$; median = 0.26 m, mean = 1.28 m, s.d. = 1.27 m; $p < 0.001$). Post hoc tests also showed that extra small (X) tools were moved significantly farther (median = 0.64 m, mean = 5.19 m, s.d. = 9.99 m) than L-tools ($p < 0.001$). Tools in the mangrove were moved farther than on the rocky shore, but this may have been due to them being easier to recover there. If stones were moved far on the rocky shore, then they were very hard to recover and probably washed away. Our control analysis of a data subset using only the first movement of each stone size in each test showed the same pattern of results, except that the S–M ($p = 0.009$) post hoc comparison did not meet the Bonferroni correction criteria for significance by a very small margin.

We also tested whether tools used for the three food groups varied in how far they were moved. We found there was a significant difference in the tool movement distance across groups (Kruskal–Wallis: $H_{3,174} = 62.93, p < 0.001$). Post hoc tests showed that tools used for sessile oysters were moved farther (median = 2.82 m, mean = 10.37 m, s.d. = 18.87 m) than for the other two groups (small gastropods: median = 0.20 m, mean = 1.77 m, s.d. = 5.09 m; large molluscs and nuts: median = 0.40 m, mean = 1.56 m, s.d. = 6.24 m; $p < 0.001$), which were pounded with stones on anvils. The farthest a tool was transported during our study was an S-sized oyster tool that was transported 99 m. The control analysis with a data subset using only the first movement of each stone size in each test showed the same pattern of results, and thus we do not report them.

(d) Stone weight at natural tool sites

We measured the mass of 375 stone tools found with food debris at naturally occurring tool-use sites. With these stones, we found the debris of 245 rock oysters (215 sessile, 30 detached), 288 nerite snails, 32 trochid snails, 18 small gastropods, 70 muricid snails, eight conches, 28 clams, 10 crabs, 138 sea almonds and five pandanus keys. The stones ranged from 16 to 5166 g (median = 282.00 g, mean = 611.85 g, s.d. = 747.72 g). We were able to categorize 318 of the stones into five groups, instead of only three, because of the large sample size. These groups were: (i) attached oysters, (ii) small gastropods (i.e. nereite and trochid snails), (iii) bivalves (i.e. clams, mussels and detached oysters), (iv) drills and conches, and (v) nuts. We then tested for variation in stone mass across groups, and found they differed significantly (Kruskal–Wallis test: $H_{4,318} = 199.73, p < 0.001$; figure 4). Mann–Whitney post hoc tests indicated each group differed significantly from each other (group 1 with 2 and 3, $p < 0.001$, and groups 2 and 3, $p = 0.007$; attached oysters: $n = 54$, median = 63 g, mean = 81.78 g, s.d. = 77.11 g, range = 16–474 g; small gastropods: $n = 139$, median = 194 g, mean = 303.60 g, s.d. = 338.18 g, range = 24–1750 g; bivalves: $n = 29$, median = 366 g, mean = 464.48 g, s.d. = 341.22, range = 90–1568 g; drill and conches: $n = 31$, median = 874 g, mean = 1118.25 g, s.d. = 778.27 g, range = 164–3286 g; nuts: $n = 65$, median = 1332 g, mean = 1590.29 g, s.d. = 959.12, range = 220–5166 g).
4. Discussion

Macaques were found to select stones that varied in mass, depending on the type of food being processed. Previous studies had shown a difference in the physical characteristics of macaque stone tools [31]. Macaques have been reported to use axe hammers, which are smaller, lighter stones used to strike open sessile oysters; and pound hammers, which are larger, heavier, harder stones used to pound open encased foods placed on rock substrates or anvil. In this study, we have confirmed this pattern and further uncovered that within the pound hammer category there is variation in mass that relates to the size and resistance of the food type. In our field experiment, we found that macaques preferred S-stones (90–110 g) for attached oysters and L-stones (400–1000 g) for small gastropods, large molluscs and nuts. However, the distribution of tool selection was more strongly biased towards L-stones for large molluscs and nuts, than for small gastropods. Furthermore, in our study on non-experimental tools collected with food debris at natural tool sites, we found we could more finely discriminate the association between tool mass and food. We showed tool mass significantly varied from smaller to larger across five food categories: (i) attached oysters, (ii) small gastropods, (iii) bivalves, (iv) drills and conches, and (v) nuts.

Our results can be compared with work on birds [13,14,37], chimpanzees [23,25] and capuchins [1–3,24,27,28] that have shown an ability to select for tool characteristics according to the task being performed. In all these species, the subjects are selecting tools with appropriate physical characteristics, such as selecting stones of proper size and mass, or sticks of proper width and length. Most relevant to our findings are field and laboratory experiments that have shown that chimpanzees and capuchins are able to select a hammer’s mass based on food type [1–3,23–25,27,28]. Our study confirms that, as in capuchins and chimpanzees, long-tailed macaques also can discriminate hammer tools of appropriate qualities, and appear to be able to select among stones that vary in the physical dimension of mass (and therefore, size) according to the foods being processed. These results provide evidence that tool-using macaques may have an ability to recognize and identify the functionality of a stone. However, even though we have identified an association between food and tool, the process by which macaques select stones and the degree of individual variation in selection preferences will require further study.

Our data collection on tool tests was constrained by environmental conditions that limited our ability to regularly access tests and to change their composition and location. We also were unable to regularly observe macaques using tests. These difficulties generated some statistical limitations, which we dealt with by analysing subsets of our data to remove potential independence errors associated with the reuse of tool tests and stones. We found that our corrected analyses did not differ very much in the pattern of results compared with our analyses of the full dataset. However, some pairwise comparisons were no longer significant after the corrected analyses. Because the general patterns were the same, we suspect the lack of significant differences in a few of the resampling methods, post hoc comparisons were only the result of sample size and loss of power, rather than biological reality. We therefore, are confident in the biological relevance of our results, despite the potential for confounds associated with independence. These issues were unavoidable during this study, but we intend to improve our ability to observe the macaques and conduct more controlled field experiments in future studies as the project progresses and we can overcome the difficulties of conducting behavioural research in an intertidal habitat.

Our study was also limited because we did not observe macaques selecting the stones we measured for this study. We were not able to see the process of selection in the tool experiments, nor could we see which macaques, or how many, were handling the stones. Do different macaques prefer different types of tools? We have seen from direct observations of macaque tool-use behaviour that a small proportion of macaques have idiosyncratic dependencies on specific tool features and reliably use the same kind of tool. However, our observations also show that a much larger proportion of macaques on Paek Nam Yai use a range of tool sizes, which seem to be in accordance with what we have found in this study (M. D. Gumert 2009–2012, unpublished data). Additionally, we did not experimentally separate the natural association between size and mass in our study, and thus cannot make conclusions on which of these two physical dimensions the macaques attended to during selection. Overall, we can conclude that macaques are able to select tools of appropriate mass for the foods they process, but, we do not yet know the mechanisms behind their selection. We anticipate future investigations observing macaque tool-use behaviour will confirm our results and provide insights into the behavioural and cognitive processes that macaques use to select stone tools for processing encased food items.

Acknowledgements. The National Research Council of Thailand provided research permits, and the National Park, Wildlife, and Plant Conservation Department of Thailand provided a park permit. We thank Low Kuan Hoong and Virnice Tan for assisting in phase 1 of this project. We greatly appreciate the constructive comments by Christian Rutz, Dora Biro, Michael Haslam and two anonymous reviewers, which helped to improve this paper. We also thank the park management at Laem Son, the local community of Bang Ben and the Andaman Peace Resort for providing assistance and support.

Funding statement. This project was funded by the Centenary Academic Development Project of Chulalongkorn University, the Academic Research Fund of the School of Humanities of Social Sciences at Nanyang Technological University, and the Ministry of Education in Singapore (grant no. RG07/95).