

# On the choice of grasp type and location when handing over an object

## Authors

F. Cini\*<sup>‡</sup>, V. Ortenzi<sup>†</sup><sup>‡</sup>, P. Corke<sup>†</sup>, M. Controzzi\*

## Affiliations

\*The Biorobotics Institute, Scuola Superiore Sant'Anna, viale Rinaldo Piaggio 34, 56025, Pisa, Italy.

<sup>†</sup>ARC Centre of Excellence for Robotic Vision, Queensland University of Technology, Brisbane, QLD, 4001, Australia.

<sup>‡</sup>These authors contributed equally to this work

## Abstract

The human hand is capable of performing countless grasps and gestures that are the basis for social activities. However, which grasps contribute the most to the manipulation skills needed during collaborative tasks, and thus which grasps should be included in a robot companion, is still an open issue. Here, we investigated the grasp choice and hand placement on objects during a handover when subsequent tasks are performed by the receiver and when in-hand and bi-manual manipulation are not allowed. Our findings suggest that in this scenario, human passers favour precision grasps during such handovers. Passers also tend to grasp the purposive part of the objects and leave "handles" unobstructed to the receivers. Intuitively, this choice allows receivers to comfortably perform subsequent tasks with the objects. In practice, many factors contribute to a choice of grasp, e.g., object constraints and task constraints. However, not all of these factors have had enough emphasis in the implementation of grasping on robots, particularly the constraints introduced by a task, which are critical to the success of a handover. Successful robotic grasping is of paramount importance if robots are to help humans with their tasks. For this reason, we believe that the results of this work can benefit the wider robotics community, as applications range from industrial cooperative manipulation to household collaborative manipulation.

## Summary

During a handover, passers prefer precision grasps and grasp purposive parts of objects leaving "handles" free for receivers.

## Introduction

Robots are increasingly asked to perform tasks in environments shared with humans and other robots, and object manipulation is a skill demanded in many of these jobs. Humans effortlessly manipulate objects and their environment, learning this skill at an impressive young age. Enriching robots with similar abilities would enable them to complete assignments in multiple situations such as households, medical care, industrial settings and agriculture. Additionally, interaction and cooperation are desirable in a plethora of circumstances, in that robots could work alongside and support workers with their workload while concurrently ameliorating the safety in such environments.

Surprisingly, robots are still far from possessing similar skills and often fail to accomplish actions as simple as grasping. Undoubtedly, a correct and stable grasp (1) is the first step towards a successful manipulation of an object. Humans grasp an abundance of objects and pay little to no attention to how they perform this action. However, grasping is hardly a simple task to execute as it heavily involves both sensory and motor control systems (2). Five major factors influence a grasp choice (3, 4): object constraints (e.g., shape, size and function); task constraints (e.g., force, and mobility); gripper constraints (e.g., the human hand or gripper kinematics, hand or gripper size relative to the object to be grasped); habits of the grasper (e.g., experience, social convention); and chance (e.g., the initial position of the object and environmental constraints). For instance, both the reaching movement of the arm and the grasping movement of the fingers may be influenced by the agents' goal and by their intention to cooperate or compete with a partner (5–7).

Intuitively, every object can be grasped in a number of ways, and the final choice of grasp type and position of the fingers on the object is dictated by a combination of the above-mentioned five factors. Generally, the emphasis is put

on a subset of these factors (3, 4, 8–10). Nonetheless, there seems to be no consensus as to which factor (if any) plays the most important role behind a grasp choice, as they possess different weight and influence in every situation. Naturally, grasping is not an action per se, but rather a purposive action (11): humans tend to grasp objects in order to use them. Hence, task and object constraints seem to be fundamental when reasoning about grasping and must be taken into due consideration when devising a grasping strategy for robots. In other words, grasp choices are strongly situation-dependant (8) and this defining characteristic justifies studies on specific situations in order to provide and gather all the information needed for a successful implementation on a robot.

On this point, robotic grasping is yet to reproduce the innate skills that humans show. Some of the reasons behind this lack of success can be traced back to difficulties in handling uncertainties in physical models, perception and control. The problem of purposive robotic grasping, while studied (12–17), has received considerably less attention than robotic object picking. Indeed, the proliferation of analytic methods (1), data-driven methods (18) and learning-based methods (19–21) has brought general advancement; however, tasks generally involve soft constraints, as is the case of bin sorting and pick-and-place tasks. For instance, pick-and-place tasks require robots to grasp and deposit objects onto surfaces or into bins, regardless of their final position and orientation. Contrarily, the use of a tool for a certain task requires robots to grip the tool stably and, contemporarily, to pick the tool so that it can be used correctly (hard constraints on the grasp).

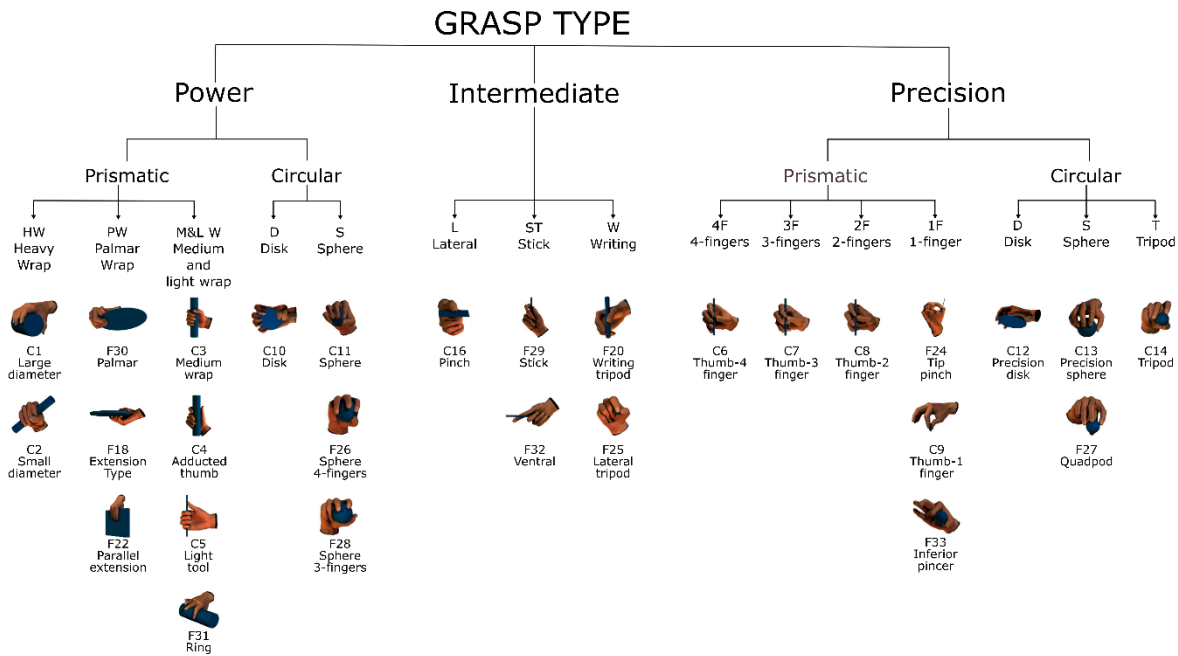
For all these reasons, manipulative skills are growingly essential in human-robot interaction. Cooperation and collaboration are complex compositions of activities that result in a natural flow of actions. Object handover is an example of joint action which involves one agent, passer, transferring an object to a partner, receiver. The implementation of an effective and natural-feeling human-robot handover is an unconquered challenge. Despite seeming a simple action, human-human handover is an effort of deduction and adjustment from both partners. Although passer and receiver share the responsibility for the stability of the object, their goal differs during the interaction (22). The passer must properly present and safely release the object to the partner, whilst the receiver's goal is to acquire a stable grasp of the object to subsequently perform tasks with it. As the actions need to be properly coordinated in time and space, passer and receiver exploit a wide range of subtle signals to predict the partner's actions and rapidly choose the correct response. For instance, gaze and the position of body and arms can be used to communicate the intent to begin the handover and predict where the handover will take place (23–26). In recent years, research on human-robot handover has focussed on improving the predictability of the robot action and the fluency of the human-robot interaction taking inspiration from human behaviour. In particular, robotic systems were developed that are able not only to control the grip force exerted on the object (27–29), but also to provide non-verbal cues similarly to humans during the interaction (25, 26, 30, 31). It was also shown that humans seldom understand the robot's actions and the interaction is perceived as unsafe or not natural when a robot does not deliver the appropriate cues (32–34). Findings of work investigating how a robot arm should be positioned to perform a handover (26, 35) showed also that the object should be presented to the human agent in its default orientations and generally positioned to allow an easy grasp by the receiver. However, these studies considered neither the passer's and receiver's grasp type nor the subsequent action the receivers have to perform with the object.

Therefore, how the selection of a grasp type and the position of the hands on the object change during the handover with respect to a non-interactive manipulation task is still a research question that needs to be answered. The present work sought to address this open issue with a view to providing useful information to design a control strategy implementable on a robotic platform for seamless and natural handover, improving the effectiveness of human-robot interaction and of robotic grasping in general. Hence in this investigation we analysed: (i) the difference between grasp types chosen for a non-interactive action and grasp types chosen for handovers; and (ii) whether a passer accounts for the receiver's task to decide how and where to grasp the objects. Considering that both hands are not always available to a human operator during some tasks and considering the challenges yet to overcome in robotic grasping and manipulation, in-hand and bi-manual manipulation have not been taken into consideration in the present work.

To this end, we conducted an experiment where 17 pairs of subjects were asked to grasp and pass 17 objects to each other. The object list includes: a closed pen (CPen), an open pen (OPen), a key (Key), a screwdriver (Screwdriver), an adversarial 3D-printed object (WShape), a plastic apple (Apple), a ball (Ball), a light disk (LDisk), a heavy disk (HDisk), a filled glass (FGlass), an empty glass (EGlass), a filled bottle (FBottle), an empty bottle (EBottle), a box of cheese crackers (Box), a thin book (Book), a metallic rod (Bar) and a toy monkey (Teddy). The experiment was composed of two sessions. In the non-interactive session (NIS), one subject (passer) was asked to grasp the objects from a table and perform two different tasks with each of the objects. In the handover session (HS), the passer was asked to hand the objects over to a partner (receiver) who subsequently performed the same two tasks the passer had performed in NIS. The first task was identical for all of the objects and consisted in putting the object onto a box. The second task was object-specific. We tracked the objects and the hands of the subjects using a vision system and manually labelled the grasp types of both agents through a frame-by-frame analysis of the video recordings. To avoid

ambiguity in the classification, we adopted the following discriminative rule. A grasp was classified as a precision grasp when only distal and intermediate phalanges were used, while the grasp was classified as a power grasp when palm and proximal phalanges were also involved. This rule prompted us to develop a taxonomy (Fig. 1), which elaborated and modified three existing taxonomies in literature (3, 8, 10). Finally, we statistically analysed the passers’ palm position relative to the objects and evaluated the differences across the four conditions of the experiment (2 sessions x 2 tasks).

Our results show that during the handover, passers prefer precision grasps, likely to maximise dexterity and leave enough space for the receiver to comfortably grasp the object. Furthermore, when in-hand and bi-manual manipulation are not allowed, passers tend to grasp the objects from the extremities and, in presence of clear affordances (i.e., portions of objects that suggest specific actions, such as handles) (36–38), the position of their hand on the objects is influenced also by the task the receiver must perform with them. This allows receivers to accept the object using immediately the most appropriate grasp to accomplish the subsequent task. This results in a high similarity between the receivers’ grasps and those exploited by subjects in NIS.

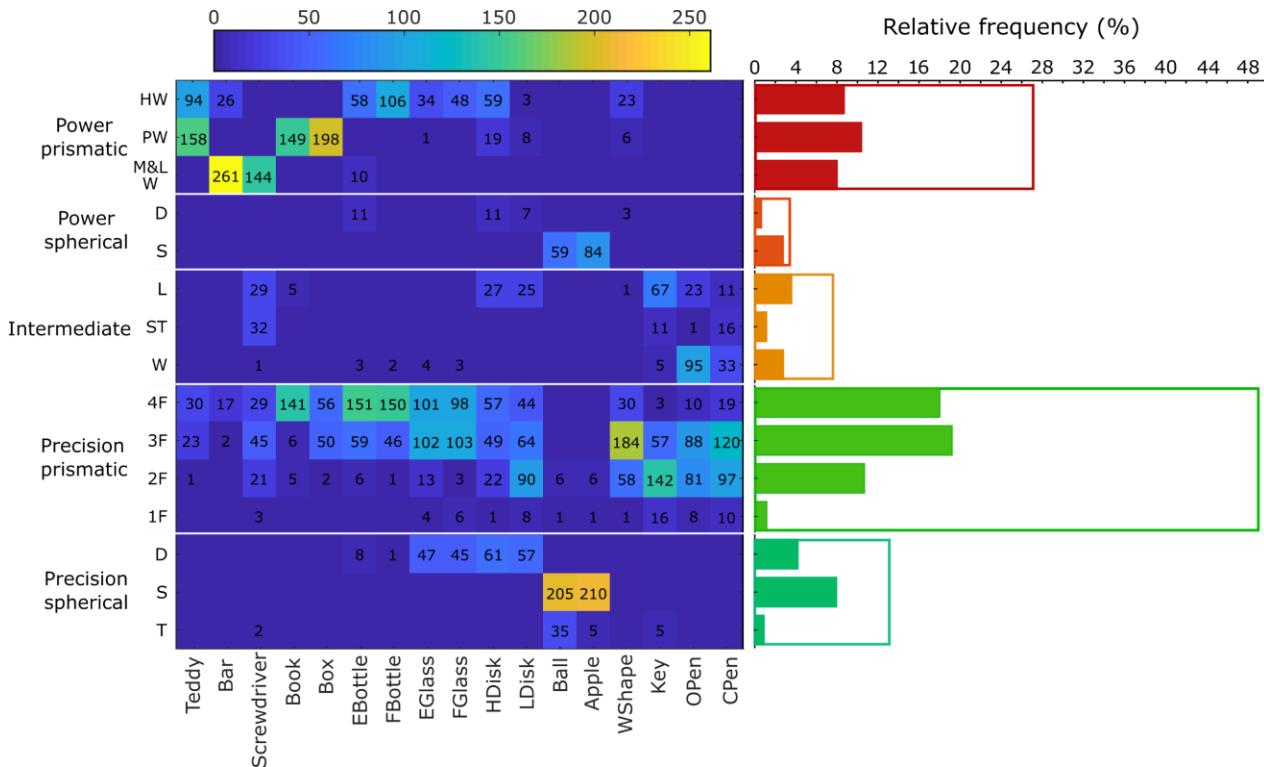


**Fig. 1 Taxonomy used to classify grasps in the experiment.** The proposed taxonomy comprises three top-level categories: power, intermediate and precision grasps. Power and precision grasps are both subdivided into circular and prismatic types. Further classifications are reported with a higher level of detail, leading to 15 classes (power prismatic: HW, PW, M&L W; power circular: D, S; intermediate: L, ST, W; precision prismatic: 4F, 3F, 2F, 1F; precision circular: D, S, T). Our analysis was focused on the above-mentioned classes. For the sake of completeness, we also reported all the 28 grasp types included in the classes and used to classify the grasps during the labelling process. For each grasp, the taxonomy reports a picture showing the grasp, and an alphabetical label in reference to the taxonomy it was taken from: C from (3), F from (8). The images are taken and adapted from (8).

## Results

### Grasp type

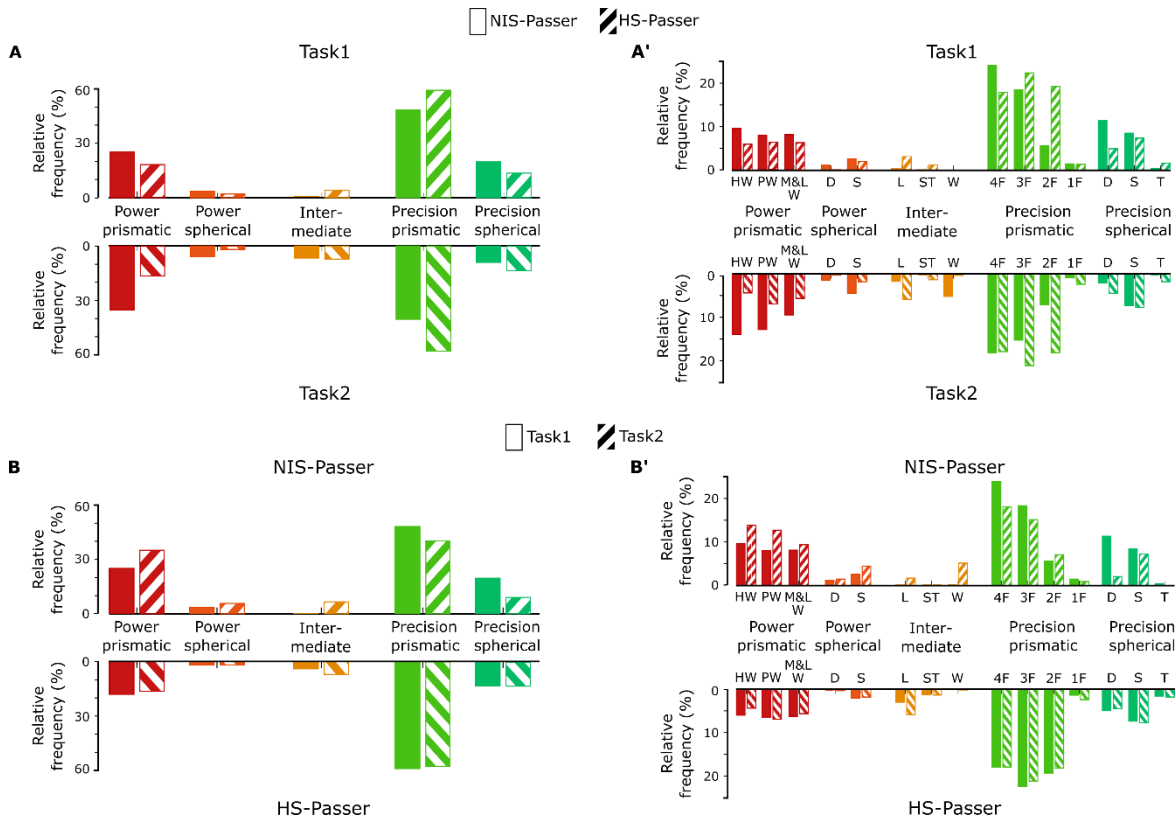
A total of 5202 grasps were labelled in this work: 1734 in NIS, 3468 in HS (1734 for the passer, 1734 for the receiver). Overall, precision, power and intermediate grasps cover respectively 62.0% (3227 occurrences), 30.4% (1580 occurrences) and 7.6% (395 occurrences) of the registered grasps (Fig. 2). Objects like Ball and Apple require a very limited set of grasps (mainly precision sphere and power sphere), while LDisk, HDisk, WShape, EBottle and Screwdriver exhibit 8 different grasps.



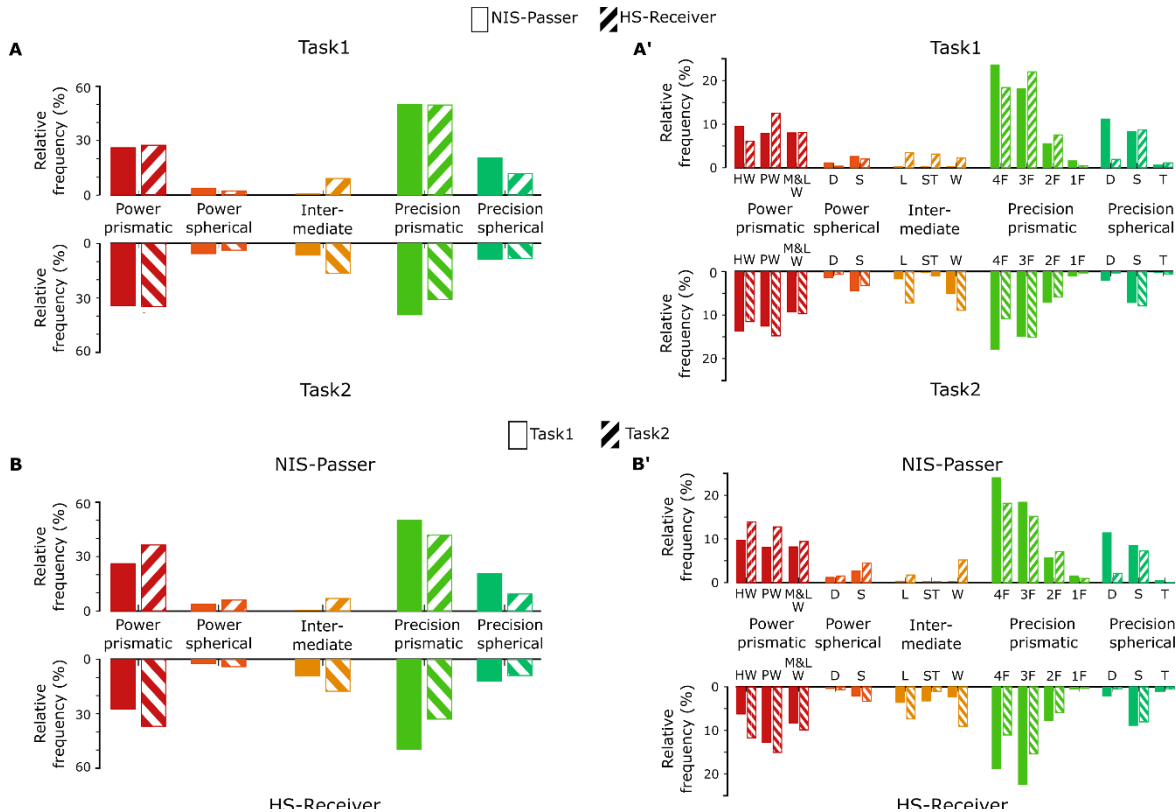
**Fig. 2 Distribution of grasps throughout the experiment.** The heat map on the left-hand side reports the occurrences of each grasp type over the 17 objects. For each object 306 grasps were labelled. The histogram on the right-hand side shows the overall frequencies of the grasp types normalized by the total number of the labelled grasps (i.e., 5202). Precision grasps are the majority (62% of all grasps), followed by 30% of power grasps and 8% of intermediate grasps.

When the performance of the passers is analysed over the tasks as in Fig. 3 (A), power grasps decrease from 29.8% in NIS to 20.9% in HS in task 1, while the drop is from 42.2% to 18.9% in task 2. Intermediate grasps are less evidently impacted by the NIS/HS condition. In particular, passers used more precision 3F and 2F and intermediate lateral in both tasks in HS than in NIS Fig. 3 (A'). When the performance of the passers is analysed over the sessions as in Fig. 3 (B), in NIS task 2 exhibits more power and intermediate grasps than task 1, with a clear increase of task-specific grasps such as Intermediate W and of Power prismatic HW and PW Fig. 3 (B'). However, the percentages of precision grasps are higher in both tasks and sessions, with a frequency of 50.9% and 73.6% respectively for NIS and HS. When comparing the grasps performed by the passers in NIS and the grasps performed by the receivers in HS, there is a certain similarity, as shown in Fig. 4. There is a noticeable equivalence between the compared tendencies, with a difference in a higher percentage of intermediate grasps in the receivers in HS.

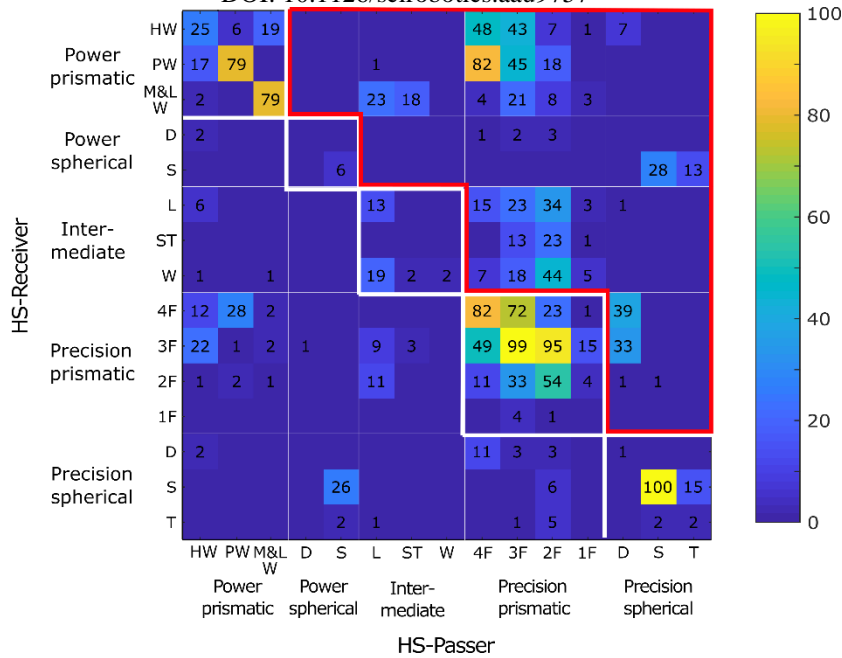
Finally, Fig. 5 shows a comparison between the grasps that passers and receivers executed during the handovers (HS). The high values in the diagonal of the heat map display a high similarity in the grasps. However, the top-right semi-plane is more populated than the bottom-left semi-plane, showing the tendency in receivers to use more power grasps than passers.



**Fig. 3. Comparison of the performances of the passers in NIS and in HS.** Histograms A show the grasp choices over the 2 tasks, while histograms B over the 2 conditions (passers in NIS and receivers in HS). Histograms A' and B' depict the same information but show higher levels of granularity in the grasp types. The frequencies shown are normalized by the total of 867 grasps performed by passers for each task of each session.



**Fig. 4. Comparison of the performances of the passers in NIS and receivers in HS.** Histograms A show the grasp choices over the 2 tasks, while histograms B over the 2 conditions (passers in NIS and receivers in HS). Histograms A' and B' depict the same information but show higher levels of granularity in the grasp types. The frequencies shown are normalized by the total of 867 grasps performed by passers or receivers for each task of each session.



**Fig. 5. Comparison of grasp types between passers and receivers during handovers.** The heatmap reports all the 1734 combinations of passer’s and receiver’s grasp observed in HS. The highest number of occurrences is found on the diagonal of this map. The red top-right semi-plane represents occurrences of more powerful grasps adopted by the receivers (with respect to passers) and is more populated than the bottom-left semi-plane (more powerful grasps by the passers with respect to receivers).

## Object occlusion

For each object and for each of the 4 conditions of the experiment (2 sessions x 2 tasks) we reported in Fig. 6 the value of Pac of each of the 17 passers, where Pac is the median value of the approaching coordinate of the passer’s palm relative to the object (defined as shown in Fig. 7) across the three repetitions of each task.

The distribution of Pac differed across objects and conditions, showing a strong influence of both task and object constraints (Fig. 6). Palm positions did not change significantly when the objects had less stringent geometric constraints, as is the case of Ball and Apple in light of their spherical symmetry. Differently, passers rearranged the position of the palm when task or object constraints were significant.

For each object, 4 comparisons (NIS task 1 vs HS task 1; NIS task 2 vs HS task 2; NIS task 1 vs NIS task 2; HS task 1 vs HS task 2) were performed using the Wilcoxon test (adjusted with the Bonferroni correction) on both the absolute value Pac ( $|Pac|$ ) and on the absolute distance between Pac and the median of the distribution of Pac of each condition ( $|dPac|$ ). In particular, we analysed  $|Pac|$  to investigate the shift of the passer’s palm position towards the extremities of the objects. Instead, we analysed  $|dPac|$  to investigate how the palm positions were clustered around the median of each condition. The complete statistical results are reported in Tab. 1.

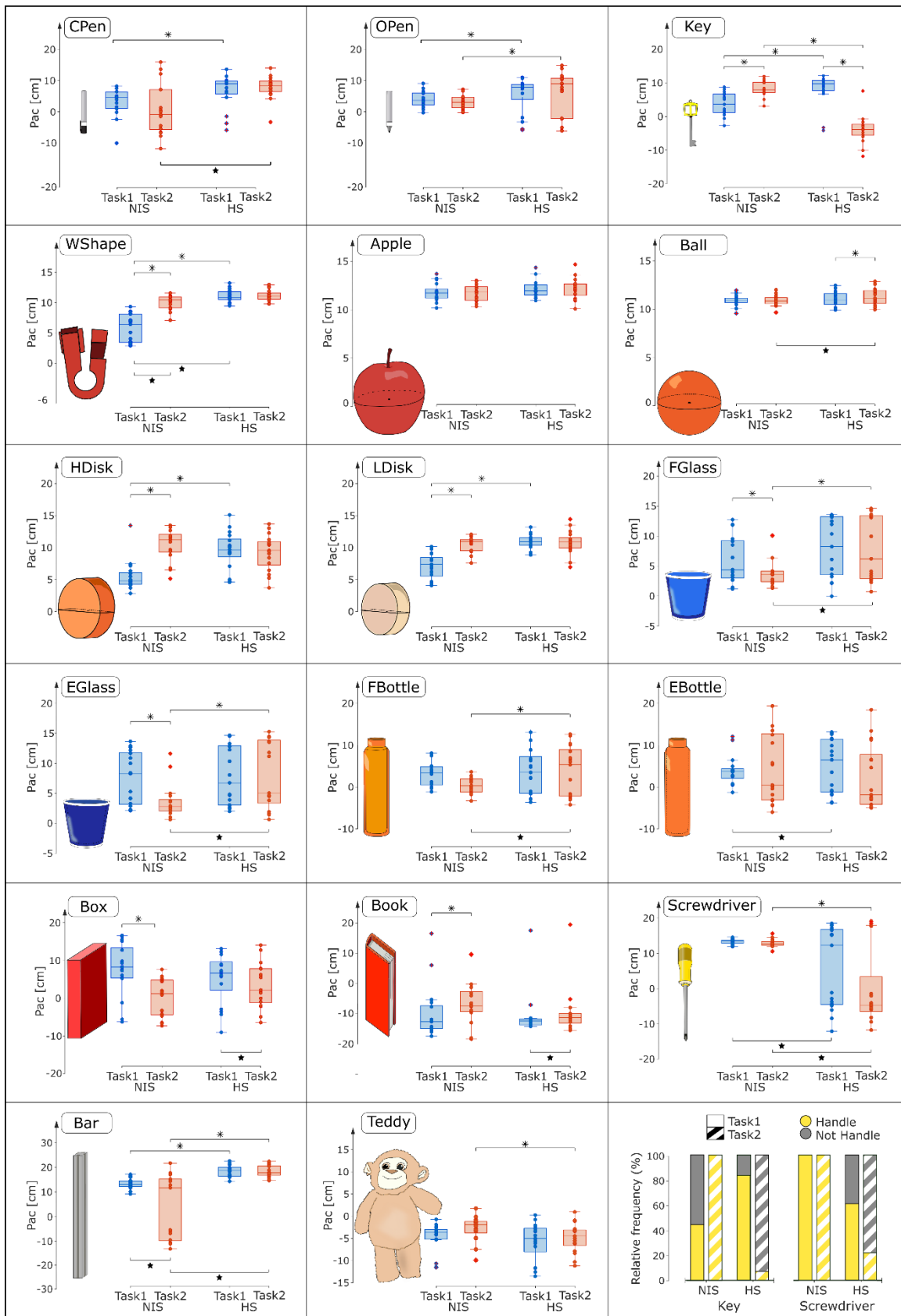
$|Pac|$  differed between NIS task 1 and HS task 1 for CPen, OPen, Key, WShape, HDisk, LDisk, Bar. Comparisons of Pac between NIS task 2 and HS task 2 were significant for OPen, Key, FGlass, EGlass, FBottle, Bar, Screwdriver and Teddy.  $|Pac|$  in NIS task 1 and in NIS task 2 were statistically different for Key, WShape, HDisk, LDisk, FGlass, EGlass, Box and Book. Finally,  $|Pac|$  significantly differs between HS task 1 and HS task 2 for the Key and the Ball. In particular, comparisons between NIS and HS show that the passer’s palm was closer to the extremities of the objects in HS than in NIS. However, when the second task was an inserting task (i.e., for the disks, EBottle, etc.), HS and NIS did not differ significantly.

The analysis on  $|dPac|$  showed that NIS task1 and HS task1 were statistically different for WShape, EBottle and Screwdriver. Interestingly  $|dPac|$  differed between NIS task 2 and HS task 2 for objects like CPen, Ball, FGlass, EGlass, FBottle, Screwdriver and Bar. Comparisons of  $|dPac|$  between NIS task 1 and NIS task 2 were significant for WShape, and Bar while HS task 1 and HS task 2 differed only for Book. These comparisons show that generally the passers’ palm positions were clustered further away from the median in HS than in NIS. However, when the second task was an inserting task, as already observed for  $|Pac|$ , HS and NIS did not differ significantly.

In session NIS, Key and Screwdriver were grasped from the handle respectively in 45.1% and in 100% of the cases in task 1. Key and Screwdriver were grasped from the handle in 100% of cases for task 2. Passers grasped Key and Screwdriver from the handle respectively in 84.3% and 60.8% of the cases in HS when the receiver had to simply put the object on the box. Interestingly, these frequencies strongly decreased when the receiver needed to grasp the handle of the object to perform the object-specific task. In this condition (HS task 2), passers grasped the Key and Screwdriver

This is the author's version of the work. It is posted here by permission of the AAAS for personal use, not for redistribution. The definitive version was published in *Science Robotics* on 13 Feb 2019: Vol. 4, Issue 27, eaau9757  
DOI: 10.1126/scirobotics.aau9757

from the handle only in 5.9% and 21.6% of cases respectively. In all the other cases the passer left the handle completely unobstructed.



**Fig. 6. Distribution of the palm position of the passers relative to the objects.** The plots show the median approaching coordinate  $Pac$  of all the 17 passers relative to each object across sessions and tasks. The significant comparisons ( $p < 0.05$ ) performed on  $|Pac|$  and on  $|dPac|$ , using the Wilcoxon test with the Bonferroni corrections, are reported with (\*) and (★) respectively. In particular,  $|Pac|$  describes how the grasping locations are shifted towards the extremities of the objects while  $|dPac|$  describes how they are clustered far from the median of the distribution. The bottom-right plot shows whether Key and Screwdriver were grasped by the handle or by the other extremity (Not Handle).



**Tab. 1 Statistical results.** Results of the comparisons performed using the Wilcoxon test on the absolute value of the passer's palm position relative to the object ( $|Pac|$ ) and the absolute distance between the median of the distribution of each condition and the passer's palm ( $|dPac|$ ). The  $p$ -values reported are adjusted according to the Bonferroni correction. Significant results ( $p < 0.05$ ) are bolded.

	NIS vs. HS				Task 1 vs. Task 2			
	TASK 1		TASK 2		NIS		HS	
	$ Pac $	$ dPac $	$ Pac $	$ dPac $	$ Pac $	$ dPac $	$ Pac $	$ dPac $
CPen	<b>Z=-2.533</b> <b>p=0.045</b>	Z=-0.544 p>0.9	Z=-1.586 p=0.451	<b>Z=-2.580</b> <b>p=0.040</b>	Z=-0.970 p>0.9	Z=-1.728 p=0.336	Z=-0.260 p>0.9	Z=-0.724 p>0.9
OPen	<b>Z=-2.959</b> <b>p=0.012</b>	Z=-0.639 p>0.9	<b>Z=-3.432</b> <b>p=0.002</b>	Z=-1.870 p=0.24	Z=-1.160 p>0.9	Z=-0.213 p>0.9	Z=-0.781 p>0.9	Z=-1.349 p>0.9
Key	<b>Z=-3.432</b> <b>p=0.002</b>	Z=-1.396 p>0.9	<b>Z=-3.053</b> <b>p=0.009</b>	Z=-0.876 p>0.9	<b>Z=-3.290</b> <b>p=0.004</b>	Z=-2.012 p=0.177	<b>Z=-3.385</b> <b>p=0.003</b>	Z=-0.260 p>0.9
WShape	<b>Z=-3.621</b> <b>p=0.001</b>	<b>Z=-2.722</b> <b>p=0.026</b>	Z=-2.438 p=0.059	Z=-0.355 p>0.9	<b>Z=-3.621</b> <b>p=0.001</b>	<b>Z=-2.689</b> <b>p=0.029</b>	Z=-0.024 p>0.9	Z=-0.260 p>0.9
Apple	Z=-2.296 p=0.087	Z=-0.402 p>0.9	Z=-2.012 p=0.177	Z=-0.639 p>0.9	Z=-0.686 p>0.9	Z=-0.166 p>0.9	Z=-0.544 p>0.9	Z=-1.302 p=0.772
Ball	Z=-1.160 p>0.9	Z=-2.012 p=0.177	Z=-2.296 p=0.087	<b>Z=-2.533</b> <b>p=0.045</b>	Z=0.213 p>0.9	Z=-0.071 p>0.9	<b>Z=-3.148</b> <b>P=0.007</b>	Z=-1.633 p=0.410
HDisk	<b>Z=-3.337</b> <b>p=0.003</b>	Z=-0.828 p>0.9	Z=-1.349 p=0.709	Z=-0.686 p>0.9	<b>Z=-3.527</b> <b>p=0.002</b>	Z=-1.065 p>0.9	Z=-0.639 p>0.9	Z=-0.308 p>0.9
LDisk	<b>Z=-3.624</b> <b>p=0.001</b>	Z=-2.059 p=0.158	Z=-0.828 p>0.9	Z=-1.396 p=0.651	<b>Z-3.574</b> <b>p=0.001</b>	Z=-1.965 p=0.199	Z=-0.923 p>0.9	Z=-1.870 p=0.246
FGlass	Z=-1.444 p=0.595	Z=-1.775 p=0.303	<b>Z=-2.533</b> <b>p=0.045</b>	<b>Z=-3.14</b> <b>p=0.007</b>	<b>Z=-2.533</b> <b>p=0.045</b>	Z=-2.391 p=0.067	Z=-0.923 p>0.9	Z=-0.592 p>0.9
EGlass	Z=-0.876 p>0.9	Z=-0.639 p>0.9	<b>Z=-2.864</b> <b>p=0.017</b>	<b>Z=-2.864</b> <b>p=0.017</b>	<b>Z=-3.527</b> <b>p=0.002</b>	Z=-2.485 p=0.052	Z=-0.497 p>0.9	Z=-0.970 p>0.9
FBottle	Z=-1.728 p=0.336	Z=-2.296 p=0.088	<b>Z=-3.574</b> <b>p=0.001</b>	<b>Z=-2.959</b> <b>p=0.012</b>	Z=-2.343 p=0.077	Z=-1.396 p=0.652	Z=-1.775 p=0.304	Z=-1.349 p=0.708
EBottle	Z=-1.775 p=0.303	<b>Z=-2.627</b> <b>p=0.034</b>	Z=-0.450 p>0.9	Z=-0.355 p>0.9	Z=-1.870 p=-0.246	Z=-2.249 p=0.098	Z=-0.118 p>0.9	Z=-0.308 p>0.9
Box	Z=-1.349 p=0.709	Z=-0.166 p>0.9	Z=-1.018 p=1.235	Z=-0.734 p>0.9	<b>Z=-3.053</b> <b>p=0.009</b>	Z=-0.544 p>0.9	Z=-2.296 p=0.0867	Z=-0.308 p>0.9
Book	Z=0.639 p>0.9	Z=-2.343 p=0.076	Z=-2.154 p=0.125	Z=-1.207 p>0.9	<b>Z=-2.675</b> <b>p=0.030</b>	Z=-0.308 p>0.9	Z=-1.491 p=0.543	<b>Z=-3.243</b> <b>p=0.005</b>
Screwdriver	Z=-1.207 p>0.9	<b>Z=-3.574</b> <b>p=0.001</b>	<b>Z=-2.533</b> <b>p=0.045</b>	<b>Z=-2.769</b> <b>p=0.023</b>	Z=-1.870 p=0.246	Z=-0.639 p>0.9	Z=-1.823 p=0.274	Z=-0.828 p>0.9
Bar	<b>Z=-3.621</b> <b>p=0.001</b>	Z=-0.718 p>0.9	<b>Z=-3.290</b> <b>p=0.004</b>	<b>Z=-3.290</b> <b>p=0.004</b>	Z=-2.107 p=0.141	<b>Z=-3.290</b> <b>p=0.004</b>	Z=0.308 p>0.9	Z=-0.118 p>0.9
Teddy	Z=-1.396 p=0.650	Z=-1.586 p=0.451	<b>Z=-2.769</b> <b>p=0.023</b>	Z=-1.065 p>0.9	Z=-1.633 p=0.410	Z=-0.024 p>0.9	Z=-1.160 p>0.9	Z=-1.633 p=0.410

## Discussion

There are multiple factors influencing a grasp, and constraints deriving from task, object and environment have to be assessed in each and every situation. Despite being a consequence of multiple sources of constraints, grasping is incontrovertibly a purposive action, and "it is the nature of the intended activity that finally influences the pattern of the grip", page 906 in (11). Anatomical considerations and constraints introduced by the objects, e.g., object size and object shape, also claim their firm role behind the decision of a grasp.

Thus, objects can generally be gripped with different grasps. However, a handover introduces additional constraints to the grasping and strongly influences the final choice. Therefore, the aim of this work was to assess the change in grasp type and in hand placement between a non-interactive action and a handover. To this end, we performed an experiment comparing subjects grasping objects for a direct use (in NIS) and for an object handover (in HS). The experimental results of this work show that passers modify their grasp choice, favouring precision grasps during the handover and

allowing the receiver to perform the same grasps used in NIS. We also observed that passers account for the receiver's subsequent task and changed their grasp strategy, especially when passers have to pass objects with a clear affordance, such as the Key and Screwdriver which are equipped with a handle.

## Ambiguities in grasp labelling

There is no doubt that grasping is a very intricate sequence of actions, and it is not surprising that its classification into a finite set of types is challenging and ambiguous. Although there seems to be some consensus on the definition of power and precision grasps (3, 4, 11, 39), every grasp aims at manipulating an object (precision element) while stably holding it (power element). Thus elements of precision handling and power grip usually coexist in the same action (11, 40, 41) and consequently, a grasp is generally classified by its predominant element.

To help disambiguate between potentially similar-looking power and precision grasps, we classified a grasp as a precision grasp when only distal and intermediate phalanges were used; whilst a grasp was classified as a power grasp when palm and proximal phalanges were also involved. This rule prompted us to partially modify the existing taxonomies in literature (3, 8, 10) in order to be as consistent as possible in the process of labelling. The discrimination between power and precision grasps performed in this work is in agreement with the previous studies about human grasp (11, 39, 42), which assess that the palm and the proximal part of the hand are involved in a power grip. Therefore, we firmly believe that it did not influence the results of the classification. Moreover, this rule implicitly suggests how much of the object surface was obstructed by the use of each grasp type. Indeed, power grasps occlude a bigger portion of the grasped object while precision grasps leave more of the surface unencumbered. This aspect is of prime importance during a handover where both the subjects have to share the limited surface of the same object. More importantly, our classification rule might be translated into a control strategy to be implemented on robots for a more purposive grasping and for a more effective handover.

## Object characteristics and task constraint do influence grasping

For our experimental setup, we used a number of objects differing in shape and size (Fig. 7). Feix et al. (43) showed that objects which have a grasp side wider than 5 cm and with mass over 250 g are more likely to be grasped with a power grasp. Counterintuitively, objects with mass less than 200 g and whose grasp side is smaller than 5 cm are not necessarily grasped using a precision grasp, as power and intermediate grasps are also frequently used. Our results show a prevalence of precision grasps (62 %) albeit the characteristics of the objects used in this experiment varied extensively, as shown in Fig. 2. Particularly, 53% of our objects had at least one dimension smaller than 5 cm and this is important since humans display a clear preference to grasp the smallest dimension of the object (43). Furthermore, the majority (59%) of the objects in our set weighed less than 200 g and only 2 objects were heavier than 500 g. Differently from the work of Feix et al. (43, 44), we observed that objects heavier than 250-300 g, such as FBottle, FGlass, HDisk and Ball, did not induce a clear preference towards power grasps during our experiment. It is our opinion that this difference may be a consequence of the different setup (environment) and tasks between the two investigations. This hypothesis is in agreement with both the discussion in Feix et al. (37), and other previous works, such as (4, 11), stating that the grasp choices can change and shift due to dissimilarities in the initial conditions, friction, personal factors (such as habits) and task demands. In (43, 44), the actions of two housekeepers and two machinists were observed during their daily work, in which the most common tasks, such as mopping or turning knobs, need some strength to be performed. On the contrary, the tasks we asked to perform were quite easy and did not require much force. Indeed, subjects in NIS and receivers in HS were first asked to simply grasp the objects and put them onto a box. Our results indicate that for this action, the object and task constraints in most of the cases were not strong enough to encourage the subjects to use power grasps rather than precision grips. Subjects in NIS and similarly receivers in HS favoured precision grasps in the 70.1% and 61.2% of cases respectively, as shown in Fig. 3 and Fig. 4. Accordingly to our hypothesis, results similar to those of Feix et al. (43, 44) have been found when the subjects were asked to use the target objects on a more specific assignment (task 2). This condition triggered the subjects in NIS and the receivers in HS to increase the frequency of power grasps up to 42.2% and 40.8% respectively. In addition, a very clear preference towards power grasps was observed for the bar (Fig. 2), the biggest object in our dataset and the only one heavier than 600g.

## Handover entails precision grasps for passer but not for receivers

A shift towards precision grasps was observed during the handover session of our experiment. In both tasks, objects were grasped preferring precision or intermediate grasps by passers in HS, differently from NIS, as shown in Fig. 3. Regardless of the task carried out by the receiver, passers exhibited a strong preference for precision grasps, recorded in 73% of the cases, whereas power grasps were displayed in less than 21% of the cases. This predisposition might be a consequence of the main objective of the passer, which is to accomplish an efficient and fluent object handover to the

receiver. To this end, it is likely that passers prefer grasp types that allow for dexterous positioning and orientation (and adjustment) of the object. In point of fact, fingertips are the area of the human hand with the highest density of mechanoreceptors (45, 46) and their use emphasises accuracy and sensitivity over stability (47, 48). In detail, the skin of the fingertips appears to have the highest density of FA I, SA I and FA II units, which convey tactile information to the brain. This is fundamental during precise/dexterous object manipulation (49), hence critical during handovers. SA I units are particularly sensitive to static forces applied on the object, therefore they play a key role on the feedback control of the grip force of the passer on the object. Differently, FA I and FA II encode fast events between the skin and the object, such as friction, and events acting on the hand-held object, such as the contact of the receiver on the object. This information is crucial for the passer to behave reactively in correction of slippage events (50) and to begin the release of the object in coordination with the grasping action of the receiver (22).

An additional point of consideration is the necessity for passers to leave enough unobstructed surface area on the object to facilitate the grasp of the receiver. This detail is in agreement with previous research in neuroscience (6, 51, 52) which shows the influence of the after-grasp task on grasp parameters such as the pre-shaping of the fingers. Unequivocally and by definition, precision grasps obstruct a smaller portion of the object than power grips, thus leaving the receivers with enough space to choose a grasp from a wider range of possibilities.

Furthermore, receivers were able to grasp and subsequently use the objects in both tasks of our experiment, exhibiting grasps similar to those employed by the subjects in session NIS, who had the only aim to use the objects directly (Fig. 3). Consequently, passers and receivers did not practise the same grasp types in HS since receivers used more power grips than passers (Fig. 5). Taken together, those outcomes suggest that unlike the passer, receivers do not shift their grasp choice toward precision during the handover since they prefer, whenever possible, a grasp that allows to accomplish the subsequent task easily and comfortably.

### Handover and subsequent action of the receiver influence the grasp location of the passer

Our experiment additionally shows that passers modulate not only their grasp type, but also their grasping location in order to allow a comfortable and ready-to-use grasp to the receiver (Fig. 6). Indeed, the analysis of the position of the palm of the passers relative to the objects shows a shift towards the extremities of the objects in HS. In the case of relatively compact objects such as WShape and the two disks, passers placed their palm further away from the centre of the object with respect to when they had to put the same objects on the box during NIS. The distribution of the approaching coordinate of the palm of the passer relative to the frame of long objects (such as the glasses, bottles, Screwdriver and the pens) tends to widen in HS, moving from the median of the distribution and clustering near the extremities of the objects.

Interestingly, comparable distributions to HS were observed in NIS when the subjects had to insert the objects into a hole (WShape, disks, EBottle, CPen, Bar). This similarity suggests that, similarly to the case of inserting tasks where subjects know that they have to leave an object side free to comfortably complete the action, passers consider the space needed by the receiver to easily grasp the object during the handover.

These results are in agreement with the qualitative study of Strabala et al.(26), suggesting that passers tend to present objects in a way that allows an easy grasping by the receiver. The same authors observed that in some instances, passers rotated objects such as a cup so that the receivers could grab the handle. Moreover, orientation is a real spatial property with distinct visual appearance that affords action (53). Our work expands those results suggesting that this behaviour is strongly influenced by the task that the receiver has to perform with the object after the handover. Indeed, we observed that when the receiver had to grasp the Key and Screwdriver by the handle to perform the object-specific task, the passer grasped the tip of the object leaving the handle free (Fig. 6). On the contrary, passers often passed the Key and the Screwdriver grasping the handle when the receiver's task was a placing task (task 1).

In a study about how humans interact and use tools, Gibson (54) defined the possibilities for action offered by objects or by the environment as object affordances. The term 'perceived affordance' was then used and diffused by Norman (55) in the context of human-machine interaction making this concept depending not only on the agents' capability but also on their goal and past experiences. In line with this theory, our results show that a clear object affordance, such as a handle, may invite different behaviours not only during single-agent actions (56, 57) but also during cooperative tasks such as handovers. In particular we suggest that, based on their past experiences, passers do reason about which area of the object can afford the receiver's subsequent task and adapt their grasp strategy to appropriately present the object to the partner.

## Conclusion and future work

Globally, our results indicate that passers do modulate the grasp not only considering their own task constraints (to fluently release the object) but also considering the requirements of the receiver's tasks in order to perform an efficient and fluent interaction. In particular, passers likely make predictions about their partner's action using the knowledge about the task that receivers want to perform after the handover and adjust their behaviour according to this prediction. This hypothesis is in agreement with a recent neuroscientific study (58) which demonstrated that when two agents are involved in a joint action requiring a degree of inter-personal coordination, they are able to predict the other agent's action (and the effects of the other agent's action) and integrate such prediction into the motor plan that guides their movements throughout the interaction (in feedforward and in feedback).

To return to our initial questions about how a handover modifies grasp choices, we suggest that there is indeed an adjustment in the passer's choice of grasp and palm placement in order to facilitate the receiver. We are fully aware that there are numerous areas in this field of research needing future work. A natural extension includes the tracking and analysis of finger placement alongside the palm. Finger placements are surrogate information for a deeper understanding of the rearrangements that passers perform during handovers. Moreover, tracking the fingers might also help with the classification of grasps, in that this data in conjunction with the disambiguation rule adopted in this paper, should decrease the ambiguity in grasp classification. On the downside, the complexity of the tracking system would increase as would the risk to affect the dexterity of the subjects influencing their grasps and their hand/fingers location on the object. In parallel, the addition of novel tasks and objects (i.e., heavier objects than those we used, made-up objects of original shapes) might bring new insights on the shift of grasp choice and on the impact that diverse tasks and object attributes have on hand placement.

In conclusion, we firmly believe that an implementation of the disambiguation rule and of the conclusions of this work will have a beneficial effect on the effectiveness of robotic grasping, impacting a comprehensive range of applications. For instance, the consideration of object and task constraints will help choosing grasp types that allow industrial robots to not only grasp objects and tools, but more importantly to use those objects and tools to successfully complete more refined tasks. In terms of collaborative manipulation, the conclusions of this work are a step forward towards augmenting the effectiveness and naturalness of these interactions as object handovers are certainly necessary for enhanced cooperation and collaboration.

## Materials and Methods

### Participants

17 pairs of subjects (for a total of 34 participants, gender: 23 males, 11 females; age: average: 30.9, standard deviation: 7.9) took part in the experiment. All participants were healthy, used their right hand and reported normal vision, and were not aware of the purpose of the experiment. All subjects participated on a voluntary basis and gave their signed consent to the participation. This project was reviewed within the Science and Engineering Faculty Low Risk Ethics process and was confirmed to be Low Risk and approved and it was confirmed to meet the requirements of the National Statement on Ethical Conduct in Human Research. This project was considered and approved by the Chair, Queensland University of Technology (QUT) Human Research Ethics Committee (UHREC) with reference number 1800000243 in the Ethics Category: Human Negligible-Low Risk.

### Experimental setup and protocol

The experimental setup consisted of a visual tracking system, 17 test-objects, 2 thin fingerless test-gloves, infrared passive markers (IRm), a box with holes of different shapes and a reference object (Fig. 7).

We used the Optitrack capture system (<http://optitrack.com/>) as visual tracking system. 10 Flex13 cameras were placed around the scene: 8 cameras were set to tracking mode and tracked the position of infrared passive markers at 120 Hz; 2 cameras were set to reference mode and recorded videos of the experiment at 30 fps (Fig. 7). The Optitrack system sent the tracking data and video recordings to a PC which ran the Motive software.

To instrument the test-objects and gloves without affecting or influencing the grasp choice of the subjects, we used different patterns and types of IRm (Fig. 7). In particular, very tiny spherical IRm (diameter of 3 mm) or IRm tape were used on little objects or whenever possible; slightly bigger IRm (diameter of 8 mm) were used only on the larger objects. The gloves were sensorised only on the back, leaving the palm surface completely free. In addition, the different patterns of IRm allowed the software to distinguish the objects, assigning a univocal coordinate frame to each of them and recording both position and orientation at 120 Hz. To simplify the data analysis, the markers were placed also so to have the axes of the objects' frames along the three major dimensions of the objects. A detailed list of the 17 test-

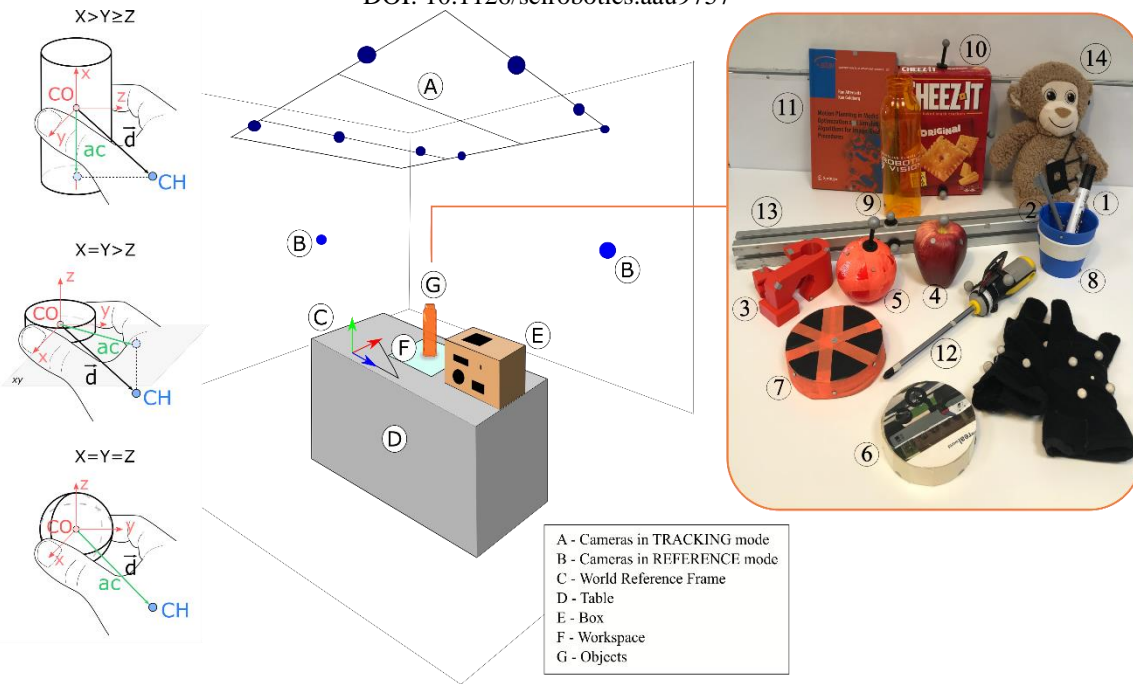
objects, describing their sizes and mass, is reported in Fig. 7. Finally, a reference object (triangular in shape and fixed on the scene) was used to set the global reference frame of the Optitrack system.

In each pair of participants, one subject was asked to play the role of passer and the other was asked to play the role of receiver. Both participants were asked to use only the right hand during the experiment and the hand was instrumented with a test-glove. The experiment comprised of two sessions. The aim of the first session (non-interactive session, NIS), was to observe how a single person grasps objects for a direct use. The second session (handover session, HS) sought to analyse the change of grasp when a subject has to pass the same objects to a partner while knowing the task the partner has to perform with it. Therefore, only the passer was involved in NIS, and the participant stood in front of a table under the visual tracking system. The subject was asked to repeatedly grasp the 17 test-objects, placed singularly on the table, and subsequently perform a task. The experiment included two tasks per object. The first task (task 1) was a general placing task and required to move the object onto a box placed on the right-hand side of the subject. The second task (task 2) was an object-specific action and differed according to the object's characteristic and function (Fig. 7). When task 2 consisted in an inserting action, subjects had to insert the object into a specific hole of the box. When task 2 was a pouring action, subjects had to pour the contents of the objects into a cup put on the table. The protocol required that the passer grasped each object three times to perform task 1 and then three times to perform task 2.

In the subsequent session, HS, passer and receiver stood (both under the visual tracking system) at opposite sides of the table and were asked to collaborate allowing the receiver to perform the same 2 tasks already carried out by the passer in NIS. In particular, the passer had to pass each of the 17 test-objects to the receiver 6 times. During the first three handovers, the receiver was asked to grasp the object from the passer and use it to perform task 1; the receiver had to perform task 2 in the following three handovers.

Participants were asked in both sessions to put the instrumented hand on the table before the beginning of each trial. Then the experimenters placed the object on the table in the appropriate position (each object had a specific initial position that was kept constant throughout the experiment in order to ensure repeatability), and informed both participants on the task to perform. A verbal command ('GO') from the experimenters signalled to start the action.

Participants were asked to use only one hand and to minimise (or avoid altogether if possible) in-hand manipulation of the objects. This instruction was motivated by the fact that: 1) considering the difficulties still present to implement robotic bi-manual and in-hand manipulation, the results of this work can still be promptly implemented on many current robotic platforms; 2) in everyday life, it is possible for receivers to ask for objects so that they are directly usable, e.g., when asking for an object without directly seeing the passers passing the object or asking for a tool while performing a job and having the other hand busy; and 3) it allowed to perform a reliable classification of the grasp used by the participants. The experimenters visually monitored the execution of the trials to verify their correctness. The trial was repeated whenever the visual tracking failed. The order of the test-objects to be grasped was chosen randomly. It is our opinion that this order was not relevant as all the objects can likely be found in all households, and they are often used in everyday life. A break of 5-10 minutes was included between the two sessions and the average duration of the experiment was 1.5 hours.



#	label	X [cm]	Y [cm]	Z [cm]	mass [g]	ac	Task1	Task2
1a	CPen	14	1.7	1.7	18	dx	Put on the box	Put into a penholder
1b	OPen	14	1.7	1.7	14	dx	Put on the box	Write
2	Key	13.5	3.7	0.8	6	dx	Put on the box	Turn into the keyhole
3	WShape	11	7	5.7	64	dx	Put on the box	Insert into a hole
4	Apple	7.7	7.7	7.7	84	$\sqrt{dx^2+dy^2+dz^2}$	Put on the box	(Pretend to) bite
5	Ball	8	8	8	270	$\sqrt{dx^2+dy^2+dz^2}$	Put on the box	(Pretend to) bite
6	LDisk	9.2	9.2	2.6	54	$\sqrt{dx^2+dy^2}$	Put on the box	Insert into a hole
7	HDisk	11.2	11.2	2.8	352	$\sqrt{dx^2+dy^2}$	Put on the box	Insert into a hole
8a	FGlass	8.3	7.8	7.8	246	dx	Put on the box	(Pretend to) drink
8b	EGlass	8.3	7.8	7.8	56	dx	Put on the box	Holding the glass with the instrumented hand and using the other hand to pretend to pour water into the glass
9a	FBottle	23	6.8	6.8	588	dx	Put on the box	Pour
9b	EBottle	23	6.8	6.8	70	dx	Put on the box	Insert into a hole
10	Box	21.5	16	6	434	dx	Put on the box	Pour
11	Book	24.2	16	1.2	362	dx	Put on the box	Put on a shelf
12	Screwdriver	27.7	3.7	3.7	166	dx	Put on the box	Screw
13	Bar	50	4	4	666	dx	Put on the box	Insert into a hole
14	Teddy	31.5	13	6	98	dx	Put on the box	Play / Shake

**Fig. 7. Experimental setup and test-objects set.** On the top left an outline of the three types of object we used, and their frame system identified by the axes  $x, y, z$  along their major dimensions  $X, Y, Z$  ( $X \geq Y \geq Z$ ) respectively. These drawings show for each type of object the distance vector  $\vec{d}$  from the centroids of the objects (CO) to the centroid of the passer's hand (CH) and the approaching coordinate of the passer's hand in the object frame (ac) are shown. On the top right, the experimental setup, the test-objects and the test-gloves are shown. The table on the bottom reports for each object: its label; its three major dimensions, its mass, the mathematical definition of ac and the tasks. We refer to dx, dy, dz as the components along  $x, y, z$  of the vector  $\vec{d}$ .

## Data analysis

We analysed the videos recorded by the 2 cameras set to recording mode and we manually classified the grasps performed by passers and receivers in each trial of NIS and HS. Illustrative examples of trials can be found in Movie M1 in the Supplementary Materials. Classifying grasps type is not trivial. This is mainly due to the fact that sometimes there are only little differences among grasps. Therefore, for labelling purposes, we used all the 28 grasp types showed in Fig. 1. Such a high level of detail in the grasp types discrimination helped us to solve ambiguities and to be consistent during the classification process.

The grasps in NIS were labelled when the passer completed the grasping action and firmly held the test-object; grasps in HS were selected as close as possible to the moment of the handover. The Motive software aligns video frames with the tracking data, hence it was possible to extract the instant of time ( $T_g$ ) at which the labelled grasps occurred through a frame-by-frame analysis. Once this analysis of the video recordings was completed, we computed the frequencies of grasp occurrences over the sessions (NIS and HS); over the objects; over the tasks (task 1 and task 2); and over the roles (passer and receiver). These analyses were performed focusing on the classes of the two lowest levels of the taxonomy tree in Fig. 1. This choice is in line with other studies in literature, as (59), where only 9 general classes of grasps are used to perform the analyses.

The position and orientation of objects and gloves worn by participants were used in order to study the obstruction of the objects' surface due to the position of the passer's hand.

We refer to the centroid of the passer hand as CH and we refer to the centroids of the objects as CO. For each object we identified its three major dimensions X, Y, Z, with  $X \geq Y \geq Z$ . The axes along these three dimensions identified the frame of each object and were denoted as x, y, z, where x was along the longest object dimension and z was along the shortest dimension respectively. The origin of the object frame was placed on CO (Fig. 7). For each time instant, we computed the homogeneous transformation matrix that related the coordinate frame of the passer's palm to the frame of the object. From the transformation matrix we obtained the distance vector  $\vec{d}$  from CO to CH with components dx, dy and dz along the axes x, y, z respectively. This allowed us to evaluate over time the approaching coordinate of the passer's hand in the object frame (ac) (Fig. 7). In particular, when  $X > Y \geq Z$ , ac was defined as the projection of  $\vec{d}$  along the major axis x and coincides with dx. When  $X = Y > Z$  (disks) the object has not a unique major axis or a defined orientation on the plan xy as x and y can be rotated round z without any change in length. Thus, in this case we defined ac as the magnitude of the projection of  $\vec{d}$  on the plane xy computed as  $\sqrt{dx^2 + dy^2}$ . When  $X = Y = Z$  (spheres), since any cross section of the object is equally circular and the object has not a major axis or a defined orientation in the 3D space, we defined ac as the magnitude of the vector  $\vec{d}$  computed as  $\sqrt{dx^2 + dy^2 + dz^2}$ . The median value of ac on the 10 samples following  $T_g$  was computed and used to calculate for each subjects the median across the three repetitions of each task of each session (Pac).

Finally, for each object and for each of the 4 conditions of the experiment (2 sessions x 2 tasks), we evaluated: the absolute value of Pac (|Pac|) and, the absolute difference between each Pac value and the median of the Pac across subjects (|dPac|). Specifically, we aimed to investigate how much the distributions of Pac were shifted towards the object's extremities and how the distributions of Pac were clustered around the median. The answers to these questions were found respectively with the analysis of |Pac| and with the analysis of |dPac| across the conditions. Therefore, we carried out 4 comparisons for each object on both |Pac| and |dPac| (NIS task 1 vs NIS task 2; HS task 1 vs HS task 2; NIS task 1 vs HS task 1; NIS task 2 vs HS task 2) using the Wilcoxon test adjusted with the Bonferroni correction. Statistical significance was defined for p-value < 0.05.

For the two test-objects with a defined handle, Key and Screwdriver, we also manually annotated how many times passers grasped them from the handle in each of the 4 conditions of the experiment.

## References and Notes

1. A. Bicchi, V. Kumar, Robotic grasping and contact: a review, in, *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation*. (IEEE, 2000; <http://ieeexplore.ieee.org/document/844081/>), vol. 1, pp. 348–353.
2. R. S. Johansson, K. J. Cole, Sensory-motor coordination during grasping and manipulative actions. *Curr. Opin. Neurobiol.* **2**, 815–823 (1992).
3. M. R. Cutkosky, On grasp choice, grasp models, and the design of hands for manufacturing tasks. *IEEE Trans. Robot. Autom.* **5**, 269–279 (1989).
4. N. Kamakura, M. Matsuo, H. Ishii, F. Mitsuboshi, Y. Miura, Patterns of static prehension in normal hands. *Am. J. Occup. Ther.* **34**, 437–445 (1980).
5. C. Ansuini, M. Santello, S. Massaccesi, U. Castiello, Effects of end-goal on Hand shaping. *J Neurophysiol.* **95**, 2456–2465 (2006).
6. C. Ansuini, L. Giosa, L. Turella, G. Altoè, U. Castiello, An object for an action, the same object for other actions: effects on hand shaping. *Exp. Brain Res.* **185**, 111–119 (2008).



7. C. Becchio, L. Sartori, M. Bulgheroni, U. Castiello, Both your intention and mine are reflected in the kinematics of my reach-to-grasp movement. *Cognition*. **106**, 894–912 (2008).
8. T. Feix, J. Romero, H.-B. Schmiemayer, A. M. Dollar, D. Kragic, The grasp taxonomy of human grasp types. *IEEE Trans. Human-Machine Syst.* **46**, 66–77 (2016).
9. I. M. Bullock, R. R. Ma, A. M. Dollar, A hand-centric classification of human and robot dexterous manipulation. *IEEE Trans. Haptics*. **6**, 129–144 (2013).
10. F. Gonzalez, F. Gosselin, W. Bachta, Analysis of hand contact areas and interaction capabilities during manipulation and exploration. *IEEE Trans. Haptics*. **7**, 415–429 (2014).
11. J. R. Napier, The prehensile movements of the human hand. *J. Bone Joint Surg. Br.* **38**, 902–913 (1956).
12. R. Detry, J. Papon, L. Matthies, Task-oriented grasping with semantic and geometric scene understanding, in, *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2017; <http://ieeexplore.ieee.org/document/8206162/>), pp. 3266–3273.
13. H. O. Song, M. Fritz, D. Goehring, T. Darrell, Learning to detect visual grasp affordance. *IEEE Trans. Autom. Sci. Eng.* **13**, 798–809 (2016).
14. D. Song, C. H. Ek, K. Huebner, D. Kragic, Multivariate discretization for bayesian network structure learning in robot grasping, in, *2011 IEEE International Conference on Robotics and Automation* (IEEE, 2011; <http://ieeexplore.ieee.org/document/5979666/>), pp. 1944–1950.
15. M. Kokic, J. A. Stork, J. A. Hausteijn, D. Kragic, Affordance detection for task-specific grasping using deep learning, in, *2017 IEEE-RAS 17th International Conference on Humanoid Robotics (Humanoids)* (IEEE, 2017; <http://ieeexplore.ieee.org/document/8239542/>), pp. 91–98.
16. M. Hjelm, C. H. Ek, R. Detry, D. Kragic, Learning human priors for task-constrained grasping, in, *Computer Vision Systems. ICVS 2015*, G. A. Nalpanidis L., Krüger V., Eklundh JO., Ed. (Springer, Cham, 2015; [http://link.springer.com/10.1007/978-3-319-20904-3\\_20](http://link.springer.com/10.1007/978-3-319-20904-3_20)), pp. 207–217.
17. A. Nguyen, D. Kanoulas, D. G. Caldwell, N. G. Tsagarakis, Detecting object affordances with convolutional neural networks, in, *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2016; <http://ieeexplore.ieee.org/document/7759429/>), pp. 2765–2770.
18. J. Bohg, A. Morales, T. Asfour, D. Kragic, Data-driven grasp synthesis—a survey. *IEEE Trans. Robot.* **30**, 289–309 (2014).
19. L. Pinto, A. Gupta, Supersizing self-supervision: learning to grasp from 50K tries and 700 robot hours, in, *Proceedings - IEEE International Conference on Robotics and Automation* (2016), vol. 2016–June, pp. 3406–3413.
20. S. Levine, P. Pastor, A. Krizhevsky, J. Ibarz, D. Quillen, Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection, in, *International Symposium on Experimental Robotics. Springer Proceedings in Advanced Robotics, vol 1.* (Springer, Cham, 2016), pp. 173–184.
21. D. Morrison, P. Corke, J. Leitner, Closing the loop for robotic grasping: a real-time, generative grasp synthesis approach, in, *Robotics: Science and Systems (RSS)* (2018).
22. A. H. Mason, C. L. MacKenzie, Grip forces when passing an object to a partner. *Exp. Brain Res.* **163**, 173–187 (2005).
23. P. Basili, M. Huber, T. Brandt, S. Hirche, S. Glasauer, Investigating human-human approach and hand-over, in, *Human Centered Robot Systems. Cognitive Systems Monographs* (Springer, Berlin, Heidelberg, Berlin, 2009), pp. 151–160.
24. M. K. Lee, J. Forlizzi, S. Kiesler, M. Cakmak, S. Srinivasa, Predictability or adaptivity?: designing robot handoffs modeled from trained dogs and people, in, *Proceedings of the 6th international conference on Human-robot interaction* (2011; <http://portal.acm.org/citation.cfm?id=1957720>), pp. 179–180.
25. K. Strabala, M. K. Lee, A. Dragan, J. Forlizzi, S. S. Srinivasa, Learning the communication of intent prior to physical collaboration, in, *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication* (2012; <http://repository.cmu.edu/robotics>), pp. 968–973.
26. K. Strabala, M. K. Lee, Dragan Anca, J. Forlizzi, S. S. Srinivasa, Towards seamless human-robot handovers. *J. Human-Robot Interact.* **2**, 112–132 (2013).
27. W. P. Chan, C. A. C. Parker, H. F. M. Van der Loos, E. A. Croft, Grip forces and load forces in handovers: implications for designing human-robot handover controllers, in, *Proceedings of the*



- seventh annual ACM/IEEE international conference on Human-Robot Interaction - HRI '12* (ACM Press, New York, New York, USA, 2012; <http://dl.acm.org/citation.cfm?doid=2157689.2157692>), pp. 9–16.
28. W. P. Chan, C. A. Parker, H. M. Van der Loos, E. A. Croft, A human-inspired object handover controller. *Int. J. Rob. Res.* **32**, 971–983 (2013).
  29. M. K. X. J. Pan, E. A. Croft, G. Niemeyer, Exploration of geometry and forces occurring within human-to-robot handovers, in *2018 IEEE Haptics Symposium (HAPTICS)* (IEEE, 2018; <https://ieeexplore.ieee.org/abstract/document/8357196/>), pp. 327–333.
  30. Aj. Moon *et al.*, Meet me where I'm gazing, in *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction - HRI '14* (ACM Press, New York, NY, 2014; <http://dl.acm.org/citation.cfm?doid=2559636.2559656>), pp. 334–341.
  31. E. C. Grigore, K. Eder, A. G. Pipe, C. Melhuish, U. Leonards, Joint action understanding improves robot-to-human object handover, in *IEEE International Conference on Intelligent Robots and Systems* (2013), pp. 4622–4629.
  32. M. Cakmak, S. S. Srinivasa, M. K. Lee, S. Kiesler, J. Forlizzi, Using spatial and temporal contrast for fluent robot-human hand-overs, in *Proceedings of the 6th international conference on Human-robot interaction - HRI '11* (ACM Press, New York, New York, USA, 2011; <http://portal.acm.org/citation.cfm?doid=1957656.1957823>), p. 489.
  33. F. Dehais, E. A. Sisbot, R. Alami, M. Causse, Physiological and subjective evaluation of a human-robot object hand-over task. *Appl. Ergon.* **42**, 785–791 (2011).
  34. K. Koay *et al.*, Exploratory study of a robot approaching a person in the context of handing over an object, in *In AAI spring symposium: multidisciplinary collaboration for socially assistive robotics (2007)*; [http://homepages.herts.ac.uk/~comrklk/pub/Koay.etal\\_AAAI\\_SS07.pdf](http://homepages.herts.ac.uk/~comrklk/pub/Koay.etal_AAAI_SS07.pdf)), pp. 18–24.
  35. M. Cakmak, S. S. Srinivasa, Min Kyung Lee, J. Forlizzi, S. Kiesler, Human preferences for robot-human hand-over configurations, in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, San Francisco, CA, USA, 2011; <http://ieeexplore.ieee.org/document/6094735/>), pp. 1986–1993.
  36. L. Montesano, M. Lopes, A. Bernardino, J. Santos-Victor, Learning object affordances: from sensory--motor coordination to imitation. *IEEE Trans. Robot.* **24**, 15–26 (2008).
  37. A. Chemero, An outline of a theory of affordances. *Ecol. Psychol.* **15**, 181–195 (2003).
  38. F. Osiurak, C. Jarry, D. Le Gall, Grasping the affordances, understanding the reasoning: toward a dialectical theory of human tool use. *Psychol. Rev.* **117**, 517–40 (2010).
  39. J. M. Landsmeer, Power grip and precision handling. *Ann. Rheum. Dis.* **21**, 164–70 (1962).
  40. T. Iberall, The nature of human prehension: Three dextrous hands in one, in *Proceedings. 1987 IEEE International Conference on Robotics and Automation* (Institute of Electrical and Electronics Engineers; <http://ieeexplore.ieee.org/document/1087813/>), vol. 4, pp. 396–401.
  41. T. Iberall, Human prehension and dexterous robot hands. *Int. J. Rob. Res.* **16**, 285–299 (1997).
  42. C. L. MacKenzie, T. Iberall, *The grasping hand* (North-Holland, 1994).
  43. T. Feix, I. M. Bullock, A. M. Dollar, Analysis of human grasping behavior: object characteristics and grasp type. *IEEE Trans. Haptics.* **7**, 311–323 (2014).
  44. T. Feix, I. M. Bullock, A. M. Dollar, Analysis of human grasping behavior: correlating tasks, objects and grasps. *IEEE Trans. Haptics.* **7**, 430–441 (2014).
  45. R. S. Johansson, Å. B. Vallbo, Tactile sensory coding in the glabrous skin of the human hand. *Trends Neurosci.* **6**, 27–32 (1983).
  46. A. B. Vallbo, R. S. Johansson, Properties of cutaneous mechanoreceptors in the human hand related to touch sensation. *Hum. Neurobiol.* **3**, 3–14 (1984).
  47. D. A. Bennett, S. A. Dalley, D. Truex, M. Goldfarb, A multigrasp hand prosthesis for providing precision and conformal grasps. *IEEE/ASME Trans. Mechatronics.* **20**, 1697–1704 (2015).
  48. M. Stachowsky, T. Hummel, M. Moussa, H. A. Abdullah, A slip detection and correction strategy for precision robot grasping. *IEEE/ASME Trans. Mechatronics.* **21**, 2214–2226 (2016).
  49. R. S. Johansson, J. R. Flanagan, Coding and use of tactile signals from the fingertips in object manipulation tasks. *Nat. Rev. Neurosci.* **10**, 345–359 (2009).
  50. R. S. Johansson, G. Westling, Roles of glabrous skin receptors and sensorimotor memory in automatic control of precision grip when lifting rougher or more slippery objects. *Exp. Brain Res.*

- 56**, 550–564 (1984).
51. C. Ansuini, M. Santello, S. Massaccesi, U. Castiello, Effects of end-goal on hand shaping. *J. Neurophysiol.* **95**, 2456–2465 (2005).
  52. J. Lukos, C. Ansuini, M. Santello, Choice of contact points during multidigit grasping: effect of predictability of object center of mass location. *J. Neurosci.* **27**, 3894–3903 (2007).
  53. E. Symes, R. Ellis, M. Tucker, Visual object affordances: object orientation. *Acta Psychol. (Amst)*. **124**, 238–255 (2007).
  54. J. J. Gibson, *The ecological approach to visual perception* (Routledge, Hillsdale, NJ., 1899; <https://www.taylorfrancis.com/books/9781315740218>).
  55. D. A. Norman, *The design of everyday things* (Originally published: The psychology of everyday things, 1988; [www.basickbooks.com](http://www.basickbooks.com)).
  56. L. Sartori, E. Straulino, U. Castiello, How objects are grasped: the interplay between affordances and end-goals. *PLoS One.* **6**, e25203 (2011).
  57. R. Withagen, H. J. de Poel, D. Araújo, G.-J. Pepping, Affordances can invite behavior: Reconsidering the relationship between affordances and agency. *New Ideas Psychol.* **30**, 250–258 (2012).
  58. L. M. Sacheli, E. Arcangeli, E. Paulesu, Evidence for a dyadic motor plan in joint action. *Sci. Rep.* **8**, 5027 (2018).
  59. M. Vergara, J. L. Sancho-Bru, V. Gracia-Ibáñez, A. Pérez-González, An introductory study of common grasps used by adults during performance of activities of daily living. *J. Hand Ther.* **27**, 225–234 (2014).

## Acknowledgments

The authors want to thank all the subjects for taking the time to participate to the experiment; and Andres Marmol and Mario Strydom for their kind help with the Optitrack motion capture device. **Funding:** V. O. and P. C. are supported by the Australian Research Council Centre of Excellence for Robotic Vision (project number: CE140100016). F.C. and M.C. are supported by the European Research Council (project acronym: MYKI, project number: 679820) **Author contributions:** F.C. and M.C. initiated this work. M.C., F.C. and V.O. designed the experiments. V.O. and F.C. built the experimental setup and performed the experiments. All authors analysed the data, discussed and wrote the manuscript. M.C., V.O. and P.C. oversaw and advised the research. P.C. provided financial support. **Competing interests:** There are no competing interests. **Data and materials availability:** Please contact F.C. for data and other materials.

## SUPPLEMENTARY MATERIALS

Movie M1. Example of trials with the views of the object tracking (the trials shown in the video were realised with the only purpose to explain the experimental procedure and were not included in the analysis of grasp type and location).