

Cognitive Science (2018) 1–39

© 2018 The Authors. Cognitive Science - A Multidisciplinary Journal published by Wiley Periodicals, Inc. on behalf of Cognitive Science Society.

ISSN: 1551-6709 online

DOI: 10.1111/cogs.12638

When Height Carries Weight: Communicating Hidden Object Properties for Joint Action

Laura Schmitz,^a Cordula Vesper,^{a,b} Natalie Sebanz,^a Günther Knoblich^a

^a*Department of Cognitive Science, Central European University*

^b*School of Communication and Culture, Aarhus University*

Received 11 May 2017; received in revised form 19 May 2019; accepted 23 May 2018

Abstract

In the absence of pre-established communicative conventions, people create novel communication systems to successfully coordinate their actions toward a joint goal. In this study, we address two types of such novel communication systems: sensorimotor communication, where the kinematics of instrumental actions are systematically modulated, versus symbolic communication. We ask which of the two systems co-actors preferentially create when aiming to communicate about hidden object properties such as weight. The results of three experiments consistently show that actors who knew the weight of an object transmitted this weight information to their uninformed co-actors by systematically modulating their instrumental actions, grasping objects of particular weights at particular heights. This preference for sensorimotor communication was reduced in a fourth experiment where co-actors could communicate with weight-related symbols. Our findings demonstrate that the use of sensorimotor communication extends beyond the communication of spatial locations to non-spatial, hidden object properties.

Keywords: Social cognition; Joint action; Coordination; Sensorimotor communication; Coordination strategy; Experimental semiotics

1. Introduction

Language plays an essential role in our lives because, among other reasons, it has the crucial function of serving as a coordination device (Clark, 1996). When we coordinate

Correspondence should be sent to Laura Schmitz, Department of Cognitive Science, Central European University, Október 6 utca 7, 1051 Budapest, Hungary. E-mail: schmitz_laura@phd.ceu.edu

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

our actions with others, verbal communication often facilitates the coordination process, helping us to achieve a joint goal (e.g., Bahrami et al., 2010; Clark & Krych, 2004; Fusaroli et al., 2012; Tylén, Weed, Wallentin, Roepstorff, & Frith, 2010). It has even been argued that it is impossible for people to engage in joint activities without communicating (Clark, 1996). For instance, if two people want to carry a heavy sofa together, they often first talk about who is going to grab which side of the sofa. Of course, communication can also be non-verbal, such as when one person points to one side of the sofa and thereby informs the other where to grab it, and the other nods in agreement.

Such non-verbal gestures as well as spoken language have a primarily communicative function; that is, they are produced to transmit information to others. However, there are also forms of communication that piggy-back on instrumental actions. By modulating movements that are instrumental for achieving a joint goal, a communicative function can be added on top. For example, when carrying a sofa together, the person walking forwards may exaggerate a movement to the left when a turn is coming up (see Vesper, Abramova, et al., 2017). Thereby, she not only performs the instrumental action of turning but also the communicative action of informing the person walking backwards about the upcoming turn. What makes these types of actions communicative is that actors systematically deviate from the most efficient way of performing the instrumental action, thereby providing additional information that enables observers to predict the actor's goals and intentions (Pezzulo, Donnarumma, & Dindo, 2013). This, in turn, facilitates interpersonal coordination and the achievement of joint goals.

Thus, “sensorimotor communication” differs from typical verbal or gestural communication in that the channel used for communication is not separated from the instrumental action (Pezzulo et al., 2013). Individuals may resort to sensorimotor communication during online social interactions where the use of language or gesture is not feasible or insufficient (Pezzulo et al., 2013)—because the verbal channel is already occupied, because a joint action requires proceeding at a fast pace that renders verbal communication impossible (Knoblich & Jordan, 2003), or because a message is cumbersome to verbalize but easy to express by a movement modulation.

To date, experimental studies (Candidi, Curioni, Donnarumma, Sacheli, & Pezzulo, 2015; Sacheli, Tidoni, Pavone, Aglioti, & Candidi, 2013; Vesper & Richardson, 2014) as well as theoretical work (Pezzulo et al., 2013) on sensorimotor communication have almost exclusively focused on joint actions where interaction partners needed to communicate about spatial locations of movement targets (but see Vesper, Schmitz, Safra, Sebanz, & Knoblich, 2016, for an exception). This raises the question of whether sensorimotor communication can more generally help with achieving joint action coordination. As there is no a priori reason to think that the flexibility of sensorimotor communication is limited, it is an open question whether its usage extends beyond the communication of spatial locations.

To address this question, the present study investigated whether sensorimotor communication provides an effective means for communicating non-spatial, hidden object properties that cannot be reliably perceived visually. In particular, our aim was to find out whether and in what way joint action partners succeed in bootstrapping a sensorimotor communication system that enables them to coordinate their actions with respect to a

hidden object property, such as when selecting objects that match in weight. If joint action partners succeed, this would provide evidence that the scope of sensorimotor communication extends from conveying information about spatial locations to non-spatial, hidden object properties. The second aim of this study was to find out whether, when given a choice, joint action partners are more likely to create sensorimotor communication systems or symbolic communication systems. Testing which system people prefer can provide insights into the driving forces behind the emergence of novel communication systems.

1.1. Previous research on sensorimotor communication

Most studies on sensorimotor communication employed interpersonal coordination tasks where task information was distributed asymmetrically such that one actor had information that the other was lacking. To provide the missing information to their co-actors, informed actors then modulated certain kinematic parameters of their actions, such as movement direction (Pezzulo & Dindo, 2011; Pezzulo et al., 2013), movement amplitude (Sacheli et al., 2013; Vesper & Richardson, 2014), or grip aperture (Candidi et al., 2015; Sacheli et al., 2013), thereby informing their co-actors about an intended goal location.

Communication can only be successful if a message is understood by the designated receiver. In particular, the extraordinary human sensitivity to subtle kinematic differences (e.g., Becchio, Sartori, Bulgheroni, & Castiello, 2008; Manera, Becchio, Cavallo, Sartori, & Castiello, 2011; Sartori, Becchio, & Castiello, 2011; Urgesi, Moro, Candidi, & Aglioti, 2006) allows observers to recognize when others' actions deviate from the most efficient trajectory. Observers rely on their own motor systems to simulate and thereby predict the unfolding of the action and the actor's goal (Aglioti, Cesari, Romani, & Urgesi, 2008; Casile & Giese, 2006; Cross, Hamilton, & Grafton, 2006; Wilson & Knoblich, 2005; Wolpert, Doya, & Kawato, 2003). When observing exaggerated actions, the observer's prediction of efficient performance is violated in a way that biases the observer to more easily detect the actor's intended movement goal. Thus, violations from efficiency facilitate the discrimination of the actor's goal while at the same time conveying the actor's communicative intent, thereby simplifying the coordination process (Pezzulo et al., 2013).

Given the focus of previous research, it seems as if the use of sensorimotor communication in joint action coordination may be restricted to a narrow domain, that is, to tasks where co-actors coordinate spatial target locations. One exception is a recent study where joint action partners exaggerated their movement amplitude to facilitate temporal coordination (Vesper et al., 2016). Thus, an important open question is whether sensorimotor communication can go beyond the exchange of spatial and temporal information. Can co-actors use sensorimotor communication to inform each other about properties that they cannot directly perceive? In some joint actions, it may be less relevant to communicate to another person where or when to grasp an object, but it may be more relevant to inform her about a property of the object itself, such as how heavy it is. For example, when one person is handing a large cardboard box over to another person, it is crucial for the latter to know how heavy the box is so that she can prepare her action accordingly.

1.2. *Communicating hidden object properties*

The weight of an object is, at least to some extent, a hidden property that cannot be reliably derived from looking at the object—just like other object properties, be it fragility, rigidity, or temperature. Although there are certain cues that help to estimate object weight, these are not always reliable and can even be misleading. For instance, there are perceptual cues to object weight such as the size of an object (e.g., Buckingham & Goodale, 2010; Gordon, Forssberg, Johansson, & Westling, 1991) and its material (e.g., Buckingham, Cant, & Goodale, 2009), and there are kinematic cues actors produce while approaching, lifting, or carrying an object (Alaerts, Swinnen, & Wenderoth, 2010; Bingham, 1987; Bosbach, Cole, Prinz, & Knoblich, 2005; Grèzes, Frith, & Passingham, 2004a,b; Hamilton, Joyce, Flanagan, Frith, & Wolpert, 2007; Runeson & Frykholm, 1983; on discrimination of movement kinematics, see Cavallo, Koul, Ansuini, Capozzi, & Becchio, 2016). However, perceptual cues such as an object's size or material can be misleading (e.g., Buckingham et al., 2009) or uninformative (e.g., for non-transparent objects such as cardboard boxes). Kinematic cues can be unreliable, or even absent, when objects are light and differ in weight only to a small extent (e.g., Bosbach et al., 2005). Thus, especially when objects do not differ in size or material, when they are relatively light, and when their weight differences are small, neither perceptual nor kinematic cues can enable two co-actors to reliably select objects of the same weight.

To investigate the emergence of non-conventional communication about hidden object properties, we therefore chose weight as an instance of an object property for which co-actors would need to bootstrap a novel communication system. Our question was whether and in what way co-actors would use sensorimotor communication to communicate about weight. Would co-actors rely on systematic modulations of instrumental actions and, if so, which kinematic parameters would they modulate?

Note that communicating about hidden object properties such as weight will necessarily differ from communicating about visually perceivable object properties such as target location. In the previously reported studies, a communicator's subtle deviations from an efficient movement trajectory were directly related to the to-be-communicated spatial property. For instance, in a study where the task-relevant parameter was grasp location, higher movement amplitude directly implied higher grasp location (Sacheli et al., 2013). In contrast, hidden properties such as weight do not share this spatial dimension with movement parameters and thus do not directly map onto spatial movement deviations—reaching for an object with a higher movement amplitude does not *in itself* imply that the object is heavy (or light). Hence, the present study required creating novel mappings between systematic movement modulations and particular weights.

What could form the basis for such novel mappings? They could be based on specific kinematic features that are normally associated with grasping objects of different weights, for example, people's tendency to grasp very heavy (and possibly slippery) objects—such as a heavy cardboard box full of books—from underneath rather than at the top. This tendency could then be exaggerated and developed into a communicative mapping between low grasps and heavy weights (and conversely between high grasps

and light weights). Another option could be to draw on metaphorical mappings, such as the basic metaphor of More is Up and Less is Down (e.g., Lakoff, 2012). This metaphorical concept could then form the basis for a communicative mapping between high grasps and heavy weights (and conversely between low grasps and light weights). Interestingly, most metaphorical concepts also have their origins in embodied experience (Lakoff, 2012). Note that metaphorical mappings are also expressed by sensorimotor means in the case of co-speech gesture, for example, when people raise their hands when speaking of a matter of high importance (on metaphorical gestures, see, for example, McNeill, 1992). In sum, a novel sensorimotor communication system for conveying hidden object properties might be based on weight-specific grasping affordances or on metaphorical associations.

Assuming that it is feasible to use sensorimotor communication to convey hidden object properties, the question emerges whether sensorimotor communication is only one *possible* way of communicating these properties or whether it is also the *preferred* way. Instead of using sensorimotor communication, co-actors might be prone to develop a novel shared symbol system by establishing arbitrary associations between particular symbols and particular instances of a hidden property, akin to natural human language (de Saussure, 1959). Thus, the second aim of this study was to investigate whether joint action partners generally prefer using sensorimotor communication or symbolic communication for communicating hidden object properties, or whether situational factors affect their preference. We designed a task that allowed for the emergence of either of these two communication systems so that we could test whether the use of sensorimotor communication extends to hidden object properties, and whether the necessity to communicate such object properties favors the emergence of a symbolic, language-like communication system.

1.3. Creating non-conventional communication systems

By addressing the emergence of novel communication systems, our study built on previous work in “experimental semiotics” (Galantucci, 2009; Galantucci, Garrod, & Roberts, 2012) which explores how novel forms of communication emerge in the laboratory when co-actors cannot rely on conventional forms of communication (de Rooter, Noordzij, Newman-Norlund, Hagoort, & Toni, 2007; Galantucci, 2005; Garrod, Fay, Lee, Oberlander, & MacLeod, 2007; Healey, Swoboda, Umata, & King, 2007; Misyak, Noguchi, & Chater, 2016; Scott-Phillips, Kirby, & Ritchie, 2009). This approach provides insights into the processes leading to the successful bootstrapping of communication systems (Galantucci et al., 2012). However, instead of analyzing the structure and development of the communication system that would evolve (e.g., Duff, Hengst, Tranel, & Cohen, 2006; Garrod, Fay, Rogers, Walker, & Swoboda, 2010; Garrod et al., 2007; Healey, Swoboda, Umata, & Katagiri, 2002; Healey et al., 2007), we focused on investigating which *type* of communication system people would choose to establish.

The task used in the present study could be solved only by creating a novel communication system in a situation where participants had nothing but their instrumental

movements to communicate. Notably, a few studies in experimental semiotics (de Ruiter et al., 2007; Scott-Phillips et al., 2009) worked with similar constraints, allowing participants to use only their instrumental actions to communicate locations in a virtual environment. However, to our knowledge, the important question of whether and how novel communication systems are established and used to communicate about the hidden properties of objects has not yet been addressed.

In the present study, we explored this question in a series of four experiments. Experiment 1 served to establish that sensorimotor communication extends to hidden properties. Building on this, Experiments 2 and 3 investigated which type of communication system participants preferentially choose to establish when given a choice. Experiment 4 asked whether the preference for sensorimotor communication observed in Experiments 2 and 3 could be reduced by providing symbols that bear an intrinsic relation to the hidden property.

2. Experiment 1

In Experiment 1, we asked whether and in what way co-actors would use sensorimotor communication to transmit information about object weight to solve a coordination problem. To address this question, we created a task where two co-actors were given the joint goal of establishing a balance on a scale. Each co-actor placed one object on one of the two scale pans. We varied whether only one co-actor (“asymmetric knowledge”) or both co-actors (“symmetric knowledge”) received information about the correct object weight beforehand.

In the asymmetric knowledge condition, the informed actor knew the correct object weight and needed to communicate this information to her uninformed co-actor to enable her to choose one of three objects of different weights to achieve the joint balancing goal. In the symmetric knowledge condition where both co-actors were informed about object weight, no information transmission was required. We predicted that co-actors would not communicate in this condition.

If informed actors engage in sensorimotor communication, they should systematically modulate kinematic parameters of their movements. In contrast to previous studies where co-actors modulated kinematic parameters that directly mapped onto the communicated spatial locations (e.g., Sacheli et al., 2013; Vesper & Richardson, 2014), co-actors in this study needed to go beyond such direct mappings and develop novel mappings. Based on our hypotheses about how such novel communicative mappings might originate (see second to last paragraph in 1.2), and based on the results of a pilot study,¹ we predicted that the most likely way to communicate object weight would be to grasp objects of different weights at different heights. Specifically, we predicted that the majority of participants in this study would choose three clearly distinct grasp heights to refer to the three different object weights. However, our pilot study had also indicated that such height-weight mappings are not the only possible way of communicating, as several participants had used alternative strategies.

Instead of using sensorimotor communication, co-actors might choose to solely rely on naturally occurring perceptual or kinematic cues to weight to achieve coordination. This is unlikely, however, because there were no visually perceivable differences between the objects used in the present experiment (same size, same material). Any natural kinematic differences were expected to be minimal because all objects were quite light and differences in weight were small. Nevertheless, to exclude the possibility that co-actors rely on natural kinematic cues, we included an individual non-communicative baseline with a separate participant sample to assess whether any weight-specific kinematic differences occur during individual reach-to-grasp actions.

2.1. Method

2.1.1. Participants

We chose a sample size of $N = 12$ (i.e., six pairs) based on the assumption that effect sizes would be quite large because for successful communication participants needed to produce modulations that could be easily identified by their task partner. Only if the communicative signal could be reliably detected despite natural variability (i.e., noise), participants would be able to create an efficient communication system. Accordingly, nine female and three male volunteers participated in randomly matched pairs in the joint condition (four only-female pairs, one only-male pair, $M_{\text{age}} = 22.7$ years, $SD = 2.95$ years, range: 19–29). The two participants in each pair did not know each other prior to the experiment. In the individual baseline, five female volunteers and one male volunteer participated individually ($M_{\text{age}} = 23.7$ years, $SD = 1.80$ years, range: 21–27).

In the joint condition, three participants were left-handed but used their right hand to perform the task. In the individual condition, all participants were right-handed. All participants had normal or corrected-to-normal vision. Participants signed prior informed consent and received monetary compensation. The study was approved by the Hungarian United Ethical Review Committee for Research in Psychology (EPKEB).

2.1.2. Apparatus and stimuli

We used two sets of three tube-like objects (height: 25.7 cm, diameter: 5.2 cm) that were visually indistinguishable. The objects were plain white (Fig. 1). They differed in weight such that there was one light (70 g), one medium (170 g), and one heavy object (270 g) in each set.

The experiment was performed using an interactive motion-capture setup (Fig. 2). Two participants were standing opposite each other at the two long sides of a table (height: 102 cm, length: 140 cm, width: 45 cm). The table was high enough such that participants could comfortably reach for objects on the table surface. A 24" Asus computer screen (resolution 1920 x 1080 pixels, refresh rate 60 Hz) was located at one short end of the table. The screen was split in half by a cardboard partition (46 cm x 73 cm) such that different displays could be shown to the two participants who could each only see one half of the screen. At each long side of the table, two circular markers (3.2 cm and 6.4 cm diameter) were located on the table surface with a distance of 45 cm between

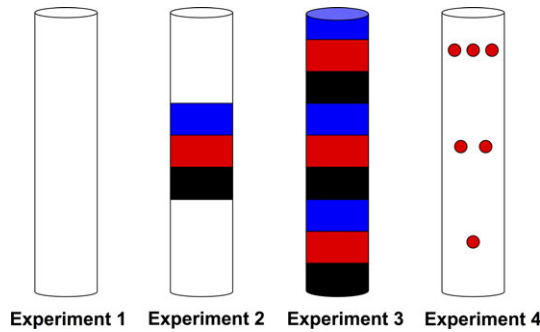


Fig. 1. The different object designs that were used in the four experiments. In each experiment, two identical sets of three objects were used, one set for each co-actor. Each object set contained one light, one medium, and one heavy object. The six objects used in each experiment were visually indistinguishable.

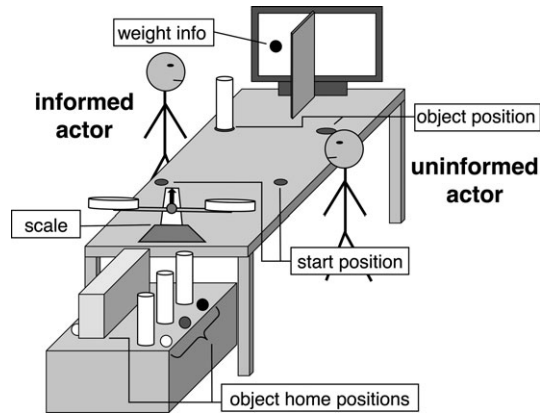


Fig. 2. Schematic depiction of the experimental setup, showing a trial from the asymmetric knowledge condition. In this condition, one object was located at the “object position” on the informed participant’s side. This participant knew the weight of the object, but her partner did not know the weight. The informed participant performed a reach-to-grasp movement from the start position toward the object and paused with her hand on the object. The uninformed participant’s task was to choose one of three objects on the home positions that matched the weight of the informed partner’s object. Once the uninformed participant had chosen an object, both participants lifted their objects and synchronously placed them on the scale.

them. The smaller circle served as the “start position” and the larger circle served as the “object position.” At the short end of the table opposite to the screen stood a mechanical scale (height: 11.5 cm, length: 26 cm, width: 11.2 cm, distance from object position: ca. 57 cm), see Fig. 2.

On a second lower table (height: 75 cm, length: 35 cm, width: 80 cm) next to the scale, the two sets of objects stood aligned before the start of each trial. These “object home positions” were marked on each of the table’s long sides. The objects were arranged in descending weight, with the heaviest object positioned closest to the participants (distances between the start position and the three object positions: 37/45/53 cm).

Next to each object position, there was a marker indicating the object's weight such that participants could easily identify the objects. This weight information was provided by circles (diameter: 5.5 cm) of three different shades such that the darker the shade, the heavier the object.² The same type of information was displayed on the computer screen to inform one or both participants (depending on the condition, see below) about the object weight in the current trial.

A Polhemus G4 electro-magnetic motion-capture system (www.polhemus.com) was used to record participants' movement data with a constant sampling rate of 120 Hz. Motion-capture sensors were attached to the nail of each participant's right index finger. Experimental procedure and data recording was controlled by a Matlab (2014a) script.

2.1.3. Procedure

Before the start of the experiment, participants were instructed in writing as well as verbally by the experimenter.

2.1.3.1. Joint condition: The two participants were informed that their task was to place objects on a scale with the joint goal to balance the scale by choosing objects of equal weight. Participants were informed that there would be three different weights. Each participant placed one object on one of the two scale pans. Participants were instructed to "coordinate and work together" to achieve the joint goal. They were not allowed to talk or gesture. Before the main experiment started, eight practice trials familiarized participants with the procedure.

In the main experiment, participants completed two blocks of 36 trials. In half of the trials, information about object weight was provided to both participants ("symmetric knowledge"), and in the other half of the trials, information was provided to only one of the participants ("asymmetric knowledge"). In the asymmetric trials, it was varied which of the two participants received information such that participants were informed equally often overall. The order of asymmetric and symmetric trials was randomized. The three different object weights occurred equally often in the two trial types in a randomized order. In total, each weight occurred 12 times in the symmetric and asymmetric knowledge conditions.

Each trial proceeded as follows (see Fig. 3): In the beginning of each trial, participants placed their index finger (with the motion sensor attached to it) on the starting position. Then, participants were verbally instructed (through a voice recording) to close their eyes. This was necessary so that participants could not observe the experimenter while she quickly arranged the objects as required for the specific trial. About 3.6 s later, participants were verbally informed who would receive weight information (e.g., "Only participant 1 receives information") and who would start the trial (e.g., "Participant 1 starts"). About 8 s after onset of the trial information, participants were verbally asked to open their eyes again ("Open!").

They could now check their side of the computer screen for weight information provided by a light, medium, or dark shaded circle (3.9 cm diameter) indicating a light, medium, or heavy object. In the symmetric knowledge condition this information was displayed on both sides of the screen; in the asymmetric knowledge condition this

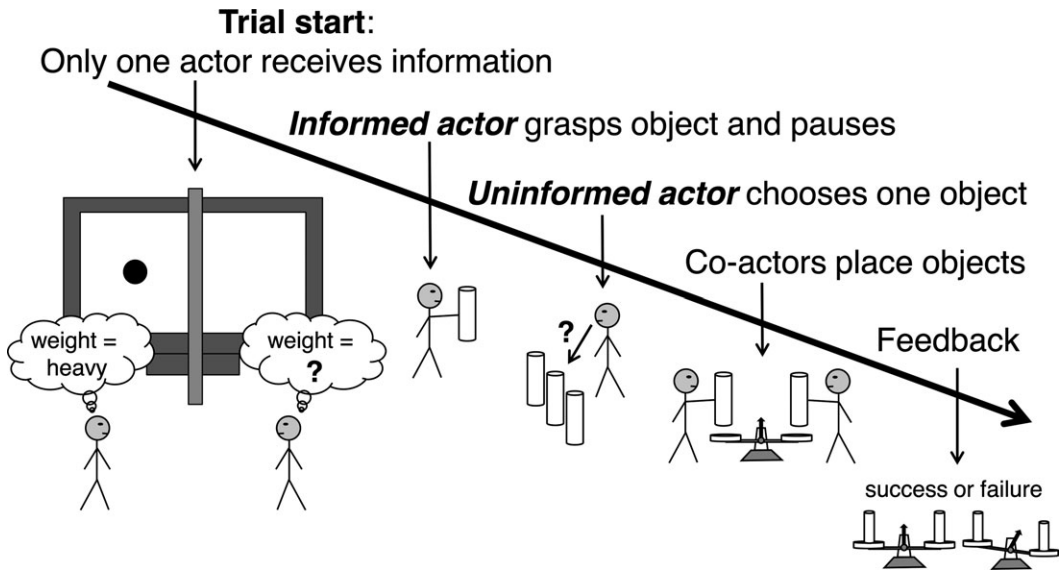


Fig. 3. Sequence of events in a trial of the joint asymmetric knowledge condition.

information was displayed only on the informed participant's side. Participants were given a time window of 3 s to process the information displayed on the screen; then they heard a short tone (440 ms, 100 Hz) that served as the start signal for the participant who initiated the joint action.

In the asymmetric knowledge condition the informed participant performed a reach-to-grasp movement toward the object positioned in front of her and paused with her hand on the object. Once the informed participant had reached the object, a short tone (660 ms, 100 Hz) was triggered to indicate that the uninformed participant could start acting. The uninformed participant's task was to choose one of three objects of different weights (located on the object home positions) that matched the weight of her partner's object. The second set of objects was covered so that the uninformed participant could not see which object was missing from her partner's object set. After the uninformed participant had chosen one of the objects, both participants lifted their objects and placed them on the respective side of the scale to receive immediate feedback about their task success. In the symmetric knowledge condition, the participant instructed to initiate the joint action reached for and grasped the object in front of her and a tone was triggered (660 ms, 100 Hz). Then, the other participant grasped the object of equal weight positioned in front of her and both participants proceeded to place the objects on the scale that always reached a balance in this condition. At the end of each trial, the experimenter removed the objects from the scale.

After the end of the experiment, participants filled out a questionnaire in which they were asked to explain how they had solved the task (i.e., whether they [and their partner] had followed a specific strategy to inform each other about the object weight).

2.1.3.2. Individual baseline: Participants were informed that their task was to lift and place objects of three different weights. They were instructed to perform their movements in a natural way. The course of each trial was kept as similar as possible to the joint condition. Participants performed 36 experimental trials (i.e., each weight occurred 12 times). Following the verbal instructions, participants closed their eyes and opened them again (just as in the joint condition, except that there was no further information given in between); then they received the weight information on the screen, heard the starting tone, and reached for and grasped the object located in front of them. After their hand had reached the object, another tone was triggered. Following the tone, participants placed the object on one side of the scale. Before the start of the experiment, participants performed three practice trials to get familiar with the procedure. After the experiment, they were asked whether they had noticed anything about the way they had performed their movements.

2.1.4. Data analysis

Data preparation was conducted in Matlab (2013b). Prior to analysis, all movement data were filtered using a fourth-order two-way low-pass Butterworth filter at a cutoff of 10 Hz. In the joint condition, the analysis of movement data focused on the participant who performed the first reach-to-grasp movement in a trial. In the individual baseline, all individual movement data from the first reach-to-grasp movement were analyzed.

Statistical analyses were performed using IBM SPSS 22 as well as customized R scripts (2016). Linear mixed models analyses were carried out using “lme4” (Bates, Maechler, Bolker, & Walker, 2015). We used linear mixed effects models to determine whether informed participants modulated their grasp height as a function of object weight. To this end, we first centered all grasping position data for each participant and trial on the mean of the medium weight. We used absolute values because we were interested in systematic differences between grasp heights for different object weights regardless of the specific mapping direction (i.e., higher grasp positions could be mapped to lighter or heavier weights). We then used the medium weight as a reference group and compared it to light weight and to heavy weight. We clustered the data by participant and by pair by modeling random intercepts. We also clustered by weight by including random slopes. In a subset of cases (i.e., eight of 16 models), random slopes could not be included because the models failed to converge. We report unstandardized coefficients, which represent the mean differences in grasp height in centimeters. Significant differences between the grasp heights for medium versus light weight and for medium versus heavy weight imply that participants consistently grasped objects of different weights at different heights. Within each experiment, we compared whether the size of the differences in grasp height between heavy and medium weight and between light and medium weight differed as a function of condition (i.e., asymmetric knowledge versus symmetric knowledge).

In addition, we derived the signed differences in grasp height between adjacent weights to determine whether participants mapped high grasps to light weights and low grasps to heavy weights, or vice versa. We computed one signed difference value per participant by first taking the difference between the mean grasp height values for medium and light

objects ($\Delta_{\text{medium-light}}$) and between the mean grasp height values for heavy and medium objects ($\Delta_{\text{heavy-medium}}$), and then averaging across these two difference values. Positive difference values imply that heavy objects were grasped at higher positions than light objects, whereas negative difference values imply the reverse.

Finally, we computed matching accuracy as a measure of joint task performance. Trials in which the two co-actors had achieved a balanced scale by choosing objects of equal weight were classified as “matching” and trials in which the co-actors had chosen objects of different weight were classified as “mismatching.” Overall accuracy was calculated as the number of matching trials as a percentage of all trials. This measure was computed only for the asymmetric condition because success was guaranteed in the symmetric condition. Statistical analyses were performed using IBM SPSS 22 as well as customized R scripts (R Core Team, 2017).

2.2. Results

Due to technical recording errors or procedural errors (e.g., participants started a trial too early), 2.3% of trials in the individual baseline and 6.3% of trials in the joint condition were excluded from the analysis.

2.2.1. Grasp height

In the joint asymmetric condition, informed participants’ grasp height for the medium weight differed significantly from the grasp height for the light weight ($B = 7.23$, $p < .001$) and from the grasp height for the heavy weight ($B = 7.83$, $p < .001$); see left panel in Fig. 4a. There was also a significant difference between the grasp heights for the medium weight and the light weight ($B = 5.55$, $p < .001$) and between the grasp heights for the medium weight and the heavy weight ($B = 6.99$, $p < .001$) in the symmetric condition (see right panel in Fig. 4a). This shows that informed actors modulated their grasp height as a function of weight regardless of whether their co-actor was informed about object weight. There was a significant difference between the size of the grasp height differences in the asymmetric and the symmetric condition for the comparison between medium and light weight ($B = 1.60$, $p = .003$), indicating that the height difference was larger in the asymmetric condition. However, there was no difference between the conditions for the comparison between medium and heavy weight ($B = 0.66$, $p = .228$).

Participants grasped light objects at the top and heavy objects at the bottom, as suggested by the negative value of the signed difference in grasp height ($M_{\text{asymmetric}} = -7.91$; $M_{\text{symmetric}} = -7.10$). Fig. 4a illustrates that in the joint condition, all participants but one mapped high grasps to light weights.

In the individual baseline, participants grasped all objects at the same height irrespective of their weight (see Fig. 4b). Participants’ grasp height for the medium weight neither differed significantly from the grasp height for the light weight ($B = 0.92$, $p = .169$) nor from the grasp height for the heavy weight ($B = 0.81$, $p = .323$).

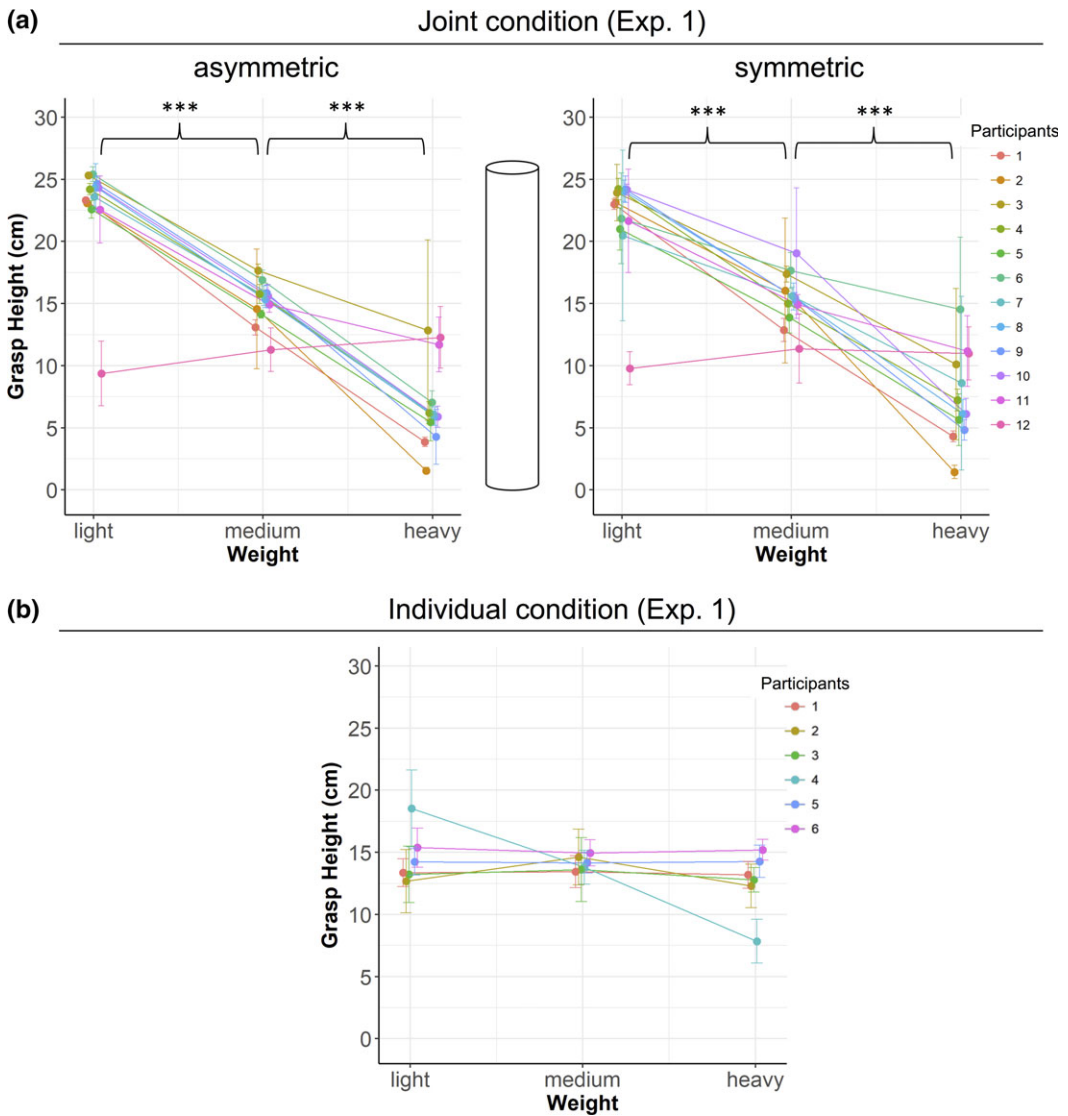


Fig. 4. Mean grasp height (a) in the joint condition and (b) in the individual baseline of Experiment 1 is shown as a function of object weight and participant. In the joint condition, all participants but one consistently grasped light objects at the top, medium objects around the middle, and heavy objects at the bottom. Participants modulated their grasp height as a function of weight regardless of whether their co-actor possessed weight information (symmetric knowledge) or not (asymmetric knowledge). In the individual baseline, participants grasped objects around the same height irrespective of their weight. The object centrally depicted in a) serves to illustrate participants' grasp height relative to the object. Error bars show standard deviations. $***p < .001$.

2.2.2. Matching accuracy

In the joint asymmetric condition, co-actors achieved an overall accuracy of 91.8% (Table 1). This value differed significantly from 33% chance performance ($t(5) = 22.85$,

Table 1

Means (and SDs) of participants' matching accuracy in the joint asymmetric condition

Experiment 1	Experiment 2	Experiment 3	Experiment 4
91.8 (9.0)	86.5 (12.3)	95.2 (7.1)	81.4 (19.7)

Note. Matching accuracy is defined as the number of trials in which co-actors chose objects of equal weight as a percentage of all trials. For an overview of trial-by-trial accuracy per participant pair, see supplementary material.

$p < .001$, Cohen's $d = 9.34$), demonstrating that informed actors successfully communicated information about the object weight to their uninformed co-actors such that these were able to choose an object of equal weight, thereby achieving the joint goal of balancing the scale.

2.2.3. Verbal reports

Eleven of 12 participants in the joint condition explicitly reported to have used height-weight mappings. Only one participant reported a different strategy, namely modulating the velocity of her reach-to-grasp movement (by moving faster for light and slower for heavy objects); see participant 12 in Fig. 4a. This participant's co-actor reported to have first used her own (height-weight) mapping and to have later adjusted to her partner's velocity modulations.

2.3. Discussion

Experiment 1 showed that informed actors transmitted information about object weight to their uninformed co-actors by systematically modulating their instrumental movements, mapping different grasp heights to different weights. This indicates that the scope of sensorimotor communication extends beyond spatial locations to hidden object properties such as weight. The behavioral findings are in line with participants' verbal reports: All participants but one reported to have used height-weight mappings.

Contrary to our prediction that actors will engage in communication only if the communicated piece of information is relevant for their co-actor (Sperber & Wilson, 1995; Wilson & Sperber, 2004), participants in Experiment 1 consistently transmitted information about the object weight, irrespective of whether their co-actor was informed or not. There may be several reasons for why informed participants engaged in communicative modulations despite the informative redundancy. First, the persistent use of communication may have served to confirm the overall functionality of the jointly established communication system and to acknowledge the joint use of the system to the co-actor, thereby demonstrating the actor's commitment to the joint action (see Michael, Sebanz, & Knoblich, 2015, 2016).

This type of behavior bears resemblance to "back-channeling" in conversation where listeners use "hmhm"-sounds, nods, or eye contact to signal their understanding to the speaker (Clark & Schaefer, 1989; Schegloff, 1968), thereby making the meaning of the speaker's utterance part of the common ground between interlocutors (Clark, 1996; Clark

& Brennan, 1991). Back-channeling is also used to show one's attentive and positive attitude toward a speaker and to indicate interest and respect for the speaker's opinion (e.g., Pasupathi, Carstensen, Levenson, & Gottman, 1999). In the present experiment, maintaining the communicative modulations despite their informative redundancy may have served a similar function as back-channeling, with the difference that it was an active (rather than responsive) means used by the communicator to demonstrate her positive and supportive attitude toward the interaction.

The redundant use of communication may have been especially critical in the very beginning of the interaction when co-actors first needed to invent and establish a communication system. By using communicative modulations when there was mutual knowledge about the weight of a given object, co-actors could rely on this common ground to make manifest the intended meaning of a specific modulation. This way, co-actors could reassure each other that they were on the same page, thereby building up new common ground and establishing a functional communication system.

A further reason for co-actors' redundant use of communication may be that trials in which both co-actors were informed randomly alternated with trials in which one of the co-actors was lacking information. Once co-actors had established a functional communication system, it may have been less costly for them to consistently adhere to the established system instead of spending extra effort to switch back and forth between a communicative and a non-communicative mode.

Experiment 1 provided initial evidence that sensorimotor communication can be used to communicate hidden object properties such as weight. In Experiment 2, we proceeded to address our second question of whether sensorimotor communication is only one *possible* way of communicating hidden object properties or whether it may also be the *preferred* way when symbolic means of communication are potentially available.

3. Experiment 2

The aim of Experiment 2 was to test which type of novel communication system two co-actors preferentially establish when faced with a coordination challenge that requires transmitting information about hidden object properties. Would they rely on sensorimotor communication, mapping particular movement deviations to particular instances of the hidden property, or would they rely on symbolic communication, establishing arbitrary mappings between particular symbols and particular instances of the hidden property?

To provide participants with an opportunity for bootstrapping a symbolic communication system, we attached patches of three different colors on each object's surface. The three different colors had no intrinsic relation with the different object weights but provided an opportunity for communicating weight in a symbolic manner by mapping specific colors to specific object weights. We predicted that actors would display this mapping by systematically grasping the object at the location of the respective color patches. However, if sensorimotor communication is preferred over symbolic forms of communication,

participants should communicate object weight by systematically grasping objects of different weights at different heights, as observed in Experiment 1.

3.1. Method

The method used in Experiment 2 was the same as in Experiment 1, with the following exceptions.

3.1.1. Participants

In the joint condition, seven female and five male volunteers participated in randomly matched pairs (two only-female pairs, one only-male pair, $M_{\text{age}} = 22.9$ years, $SD = 2.69$ years, range: 19–30). In the individual baseline, three female and three male volunteers participated individually ($M_{\text{age}} = 23.0$ years, $SD = 1.29$ years, range: 21–25). All participants were right-handed.

3.1.2. Apparatus and stimuli

The objects were white, with three colored stripes (width: ~ 3 cm) taped horizontally around each object's midsection (i.e., from 8 up to 17 cm), see Fig. 1. The three stripes were colored in blue, red, and black, respectively (from top to bottom).

3.1.3. Data analysis

To assess whether participants created stable mappings between colors and object weights, we analyzed whether participants grasped the objects at the different color patches. As we could not exclude the possibility that participants would use not only the three colors of the patches but also the white spaces of each object, we included a white category into our analysis. For each weight, we computed a rank order of colors depending on the frequency with which a participant touched the four different colors for each object weight. We then applied a sampling-without-replacement procedure mapping the most frequently touched colors to object weights, making sure that each color is only selected once (see Fig. S1 in supplementary material for an exemplification).

This procedure gave us one value (in %) per weight per participant, indicating how often each participant had used the selected color in those trials in which the given weight had occurred. If a perfectly consistent color-weight mapping was applied for all weights in all trials, the three “color usage percentages” should all be significantly above chance level. If no consistent color-weight mapping was applied, these values should not differ from chance. For all three weights, the “color usage percentages” were tested against the chance level of 25% using one-sample t tests. A significant difference in all tests implies that participants consistently grasped objects of different weights at different color patches.

3.2. Results

Due to technical recording errors or procedural errors, 2.8% of trials in the individual baseline and 4.2% of trials in the joint condition were excluded from the analysis.

3.2.1. Grasp height

In the joint asymmetric condition, informed participants' grasp height for the medium weight differed significantly from the grasp height for the light weight ($B = 4.02$, $p = .034$) and from the grasp height for the heavy weight ($B = 3.84$, $p = .032$); see left panel in Fig. 5a. There was also a significant difference between the grasp heights for the medium weight and the light weight ($B = 3.53$, $p < .001$) and between the grasp heights for the medium weight and the heavy weight ($B = 3.67$, $p < .001$) in the symmetric condition (see right panel in Fig. 5a). There was no significant difference between the size of the grasp height differences in the asymmetric and the symmetric condition (medium vs. light: $B = 0.58$, $p = .288$; medium vs. heavy: $B = 0.25$, $p = .648$). As in Experiment 1, these results indicate that informed actors modulated their grasp height as a function of weight, regardless of whether their co-actor possessed weight information or not. Also in line with Experiment 1, light objects were predominantly grasped at the top and heavy objects at the bottom, as suggested by the negative value of the signed difference in grasp height ($M_{\text{asymmetric}} = -1.65$; $M_{\text{symmetric}} = -1.76$).

Fig. 5a illustrates the inter-individual differences between participants in the joint condition. Half of the pairs (i.e., three of six) mapped specific grasp heights to specific weights. Two of these pairs (participant numbers 1–4) mapped high grasps to light weights, whereas one pair (participant numbers 11–12) used the reverse mapping. The other three pairs did not show any weight-specific differences in grasp height (see section 3.2.4). Interestingly, participant 11 modulated her grasp height only in the asymmetric condition when her partner was uninformed about object weight but not in the symmetric condition when her partner was informed. All other participants who used grasp height differences to communicate weight did so irrespectively of whether their partner was informed or not.

As in Experiment 1, participants' grasp height in the individual baseline did not differ for different weights (see Fig. 5c). Grasp height for the medium weight neither differed significantly from the grasp height for the light weight ($B = 0.15$, $p = .243$) nor from the grasp height for the heavy weight ($B = 0.20$, $p = .312$).

3.2.2. Color choice

For all weights, the “color usage percentages” were tested against the chance level of 25% using one-sample t tests to determine whether participants had consistently grasped objects of different weights at different color patches (in which case the three “color usage percentages” should all be significantly above chance level). The analysis showed that informed actors in the joint condition did not establish a consistent mapping between the colors and the three object weights (see Fig. 5b for descriptive results). In the asymmetric condition, only one of the three one-sample t tests for the three weights reached significance, given the Bonferroni-corrected significance level of $p = .008$ accounting for multiple comparisons (light: $t(11) = 1.48$, $p = .166$, Cohen's $d = 0.43$; medium: $t(11) = 5.77$, $p < .001$, Cohen's $d = 1.67$; heavy: $t(11) = 1.60$, $p = .137$, Cohen's $d = 0.46$). The same was true for the symmetric condition (light: $t(11) = 1.27$, $p = .230$, Cohen's $d = 0.37$; medium: $t(11) = 4.70$, $p = .001$, Cohen's $d = 1.36$; heavy:

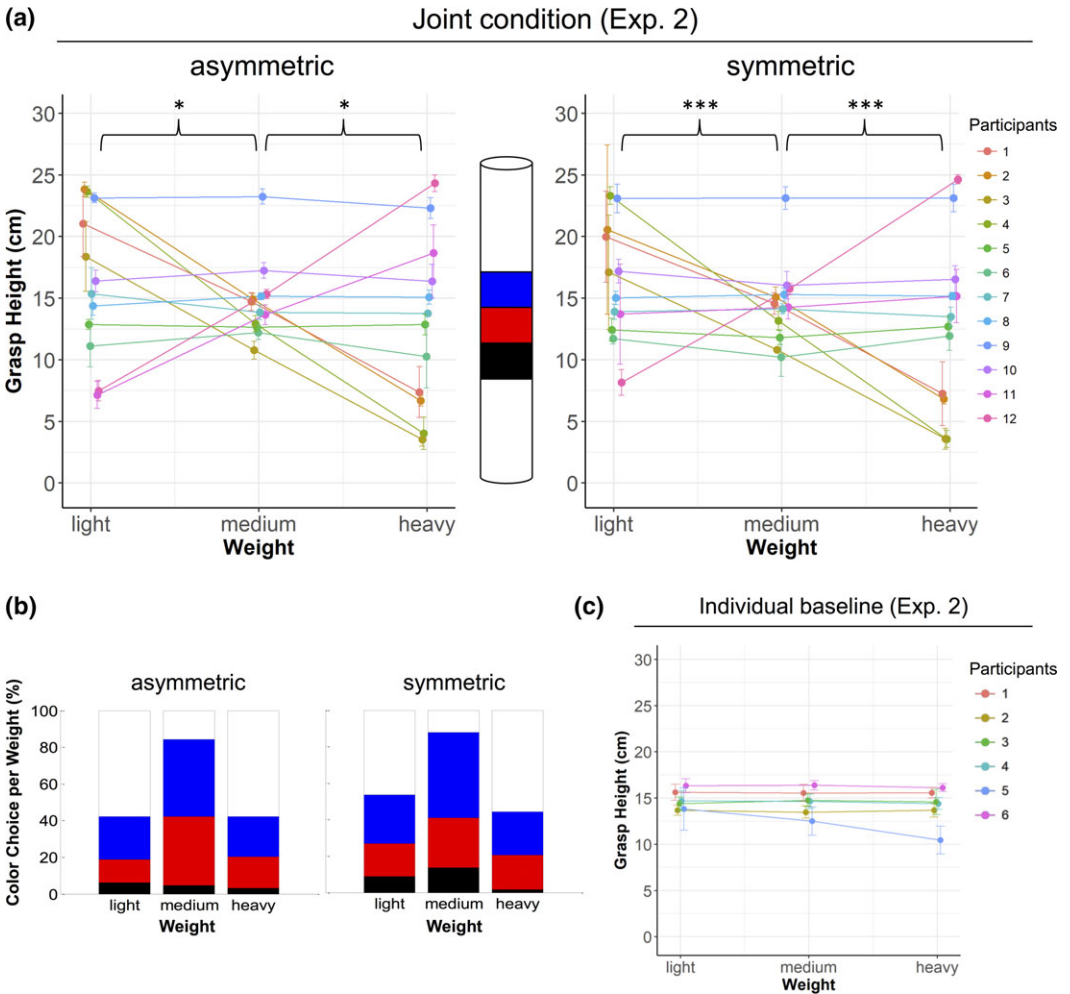


Fig. 5. Mean grasp height (a) in the joint condition and (c) in the individual baseline of Experiment 2 is shown as a function of object weight and participant. In the joint condition, six of 12 participants systematically modulated the height of their grasp as a function of object weight. Grasp height was not modulated in the individual baseline. The object centrally depicted in (a) illustrates where the three colored stripes were located relative to participants' grasp positions. Error bars show standard deviations. Panel (b) shows mean color choice (in % of trials) per weight and condition across participants. * $p < .05$, *** $p < .001$.

$t(11) = 2.03, p = .067, \text{Cohen's } d = 0.59$). These results imply that only when grasping the medium weight, participants used a particular color with a frequency above chance.

3.2.3. Matching accuracy

Co-actors were successful in reaching the joint goal, as shown by an accuracy of 86.5%. This value differed significantly from chance performance (33%), $t(5) = 15.71, p < .001, \text{Cohen's } d = 6.41$.

3.2.4. Verbal reports

Half of the participants (i.e., three of six pairs) in the joint condition explicitly reported to have used a height-weight mapping to communicate the object weight to their partner (see participant numbers 1–4 and 11–12 in Fig. 5a). Two of these pairs mapped high grasps to light objects and low grasps to heavy objects; one pair used the reverse mapping. Two pairs reported to have modulated the velocity of their reach-to-grasp movements to communicate object weight (by moving faster for light and slower for heavy objects); see participants 7–10. One pair modulated the height of the reach-to-grasp trajectory but not the endpoint of the trajectory (i.e., the grasp height at the object); see participants 5–6.

3.3. Discussion

Experiment 2 provided first evidence that actors prefer to transmit information about object weight to their co-actors by modulating their instrumental actions rather than by establishing a mapping between particular colors and particular weights. The majority of participants systematically mapped particular grasp heights to particular weights, whereas a minority of participants used alternative strategies such as modulating their movement velocity or the amplitude of their reach-to-grasp movement. Thus, participants seem to have preferred sensorimotor communication, where they communicatively modulated the instrumental action of grasping the object, over symbolic communication that would have involved a systematic grasping of the weight-unrelated color stripes on the objects. These behavioral findings are supported by participants' verbal reports as using height-weight mappings was the most frequently reported strategy.

The results from the color analysis showed that participants used a particular color with a frequency above chance only when grasping the medium weight. However, using a particular color for just one weight is not sufficient to discriminate between three different weights and thus does not establish an effective communication system. For communication to be efficient and reliable, three consistent color-weight mappings would be required. The reason for participants' consistent grasp of a specific color for the medium weight is that all of the different communication systems implied grasping the object around its middle for medium weights, resulting in a consistent grasping of the color blue for medium weights across participants.

4. Experiment 3

Based on the findings from Experiment 2, one cannot yet conclude that people *generally* prefer to communicate object weight by relying on sensorimotor communication instead of developing a symbolic communication system based on color-weight mappings. In fact, the observed preference may be due to aspects of the task design. Participants may have chosen not to rely on color-weight mappings because the color stripes had been

taped adjacently around the objects' midsections, thus requiring quite close attention from observers to discriminate the color a particular grasp was aimed at. In contrast, using the large-scale grasp height differences may have provided a less ambiguous and more obvious way of communicating.

To determine whether this was the reason that prevented participants from establishing color-weight mappings, we changed the object design in Experiment 3. We attached a multitude of colored stripes (i.e., 3×3 different colors) such that the whole object was covered from bottom to top, thereby allowing for a more distinct and large-scale grasping at specific color regions (see Fig. 1). In this way, we kept the color design used in Experiment 2 but avoided the potential problem of the stripes' close adjacency.

A further reason for the new design in Experiment 3 was that the new color configuration even allowed for a redundant use of color *and* grasp height, as participants may, for instance, grasp light objects at a "high red" position and heavy objects at a "low black" position. If participants disregarded the colors in Experiment 2 because of the proximity of different color patches, then participants in Experiment 3 would be expected to be more likely to (also) use color-weight mappings. If participants have a general preference for sensorimotor communication, then they should again use grasp height and disregard the opportunity to establish color-weight mappings.

4.1. Method

The method used in Experiment 3 was the same as in previous experiments, with the following exceptions.

4.1.1. Participants

In the joint condition, six female and six male volunteers participated in randomly matched pairs (two only-female pairs, two only-male pairs, $M_{\text{age}} = 21$ years, $SD = 1.41$ years, range: 18–23). One participant was left-handed but used his right hand to perform the task. In the individual baseline, four female and two male volunteers participated individually ($M_{\text{age}} = 23.2$ years, $SD = 1.34$ years, range: 22–25).

4.1.2. Apparatus and stimuli

Nine colored stripes were taped horizontally around each object (stripe width: ~ 2.9 cm); see Fig. 1. Colors alternated in the same way as in Experiment 2; that is, the stripes were colored in blue, red, and black (from top to bottom), with this alternation repeating three times.

4.1.3. Data analysis

For the color analysis, the "color usage percentages" were tested against the chance level of 33% (instead of 25% as in Experiment 2), because there were only three color choices available in Experiment 3.

4.2. Results

Due to technical recording errors or procedural errors, 6.9% of trials in the individual baseline and 6% of trials in the joint condition were excluded from the analysis.

4.2.1. Grasp height

In the joint asymmetric condition, informed participants' grasp height for the medium weight differed significantly from the grasp height for the light weight ($B = 7.88$, $p < .001$) and from the grasp height for the heavy weight ($B = 8.51$, $p < .001$); see left panel in Fig. 6a. There was also a significant difference between the grasp heights for the medium weight and the light weight ($B = 7.33$, $p < .001$) and between the grasp heights for the medium weight and the heavy weight ($B = 8.00$, $p < .001$) in the symmetric condition (see right panel in Fig. 6a). There was no significant difference between the size of the grasp height differences in the asymmetric and the symmetric condition (medium vs. light: $B = 0.52$, $p = .208$; medium vs. heavy: $B = 0.48$, $p = .251$).

Participants on average grasped light objects at the top and heavy objects at the bottom, as suggested by the negative value of the signed difference in grasp height ($M_{\text{asymmetric}} = -3.95$; $M_{\text{symmetric}} = -3.29$). Fig. 6a illustrates that all of the participants in the joint condition mapped specific grasp heights to specific object weights. The majority of four of six pairs mapped high grasps to light objects; only two pairs used the reverse mapping.

In the individual baseline, participants grasped objects at the same height irrespective of their weight (see Fig. 6c). Participants' grasp height for the medium weight neither differed significantly from the grasp height for the light weight ($B = -0.13$, $p = .369$) nor from the grasp height for the heavy weight ($B = 0.01$, $p = .948$).

4.2.2. Color choice

The analysis of color choice showed no consistent mappings between colors and object weights, neither in the asymmetric nor in the symmetric condition (see Fig. 6b for descriptive results). In both conditions, none of the three-one-sample t tests for the three weights reached significance, given the Bonferroni-corrected significance level of $p = .008$ accounting for multiple comparisons (asymmetric: light: $t(11) = 1.56$, $p = .148$, Cohen's $d = 0.49$; medium: $t(11) = 2.05$, $p = .065$, Cohen's $d = 0.59$; heavy: $t(11) = 0.28$, $p = .788$, Cohen's $d = 0.08$; symmetric: light: $t(11) = 1.16$, $p = .272$, Cohen's $d = 0.33$; medium: $t(11) = 0.68$, $p = .508$, Cohen's $d = 0.20$; heavy: $t(11) = 2.69$, $p = .021$, Cohen's $d = 0.78$).

4.2.3. Matching accuracy

Co-actors were successful in reaching the joint goal, as shown by an accuracy of 95.2%, which differed significantly from 33% chance performance, $t(5) = 29.77$, $p < .001$, Cohen's $d = 12.16$.

4.2. Discussion

Experiment 3 corroborated the findings from Experiment 2, showing that co-actors systematically modulated their grasp height to communicate object weight. As in Experiment 2, sensorimotor communication was preferred over symbolic color-weight mappings. Thus, it is unlikely that participants' disregard of color had been caused by the specific color arrangement in Experiment 2. These behavioral findings are in line with participants' verbal reports as all participants in the joint condition reported to have used a height-weight mapping to communicate the object weight to their partner.

Notably, the reason why participants in Experiment 3 predominantly grasped the objects at positions colored in blue (see Fig. 6b) is that the blue color stripes coincided with convenient positions for low, medium, and high grasps (see Fig. 6a). Thus, participants who used a height-weight mapping coincidentally touched the blue color stripes. This does not indicate a color-coding, however, as they touched the blue color irrespective of object weight. One might even speculate that some participants might have deliberately chosen to use the same color for all weights throughout to avoid potential confusion about whether color or grasp height was the crucial signal dimension.

5. Experiment 4

Experiments 1–3 consistently showed that co-actors choose to communicate the hidden object property weight by systematically modulating their instrumental movements, even when given the choice of creating a communication system based on color-weight mappings. On the basis of this finding, we asked whether the reason co-actors preferred sensorimotor communication was that color does not bear any intrinsic relation to weight. Any color-weight mappings would have been arbitrary and would have required participants to establish a mapping by trial and error, relying on feedback about whether or not the intended meaning and the interpretation of a symbol were correctly matched.

More specifically, what color—in contrast to grasp height—is lacking is an ordinal structure that can be systematically mapped to the ordered series of weights.³ Thus, it might not have been the arbitrariness of mapping color to weight, but rather the absence of ordinal relations between the colors matching the ordinal relations between the weights, that prevented participants from creating color-weight mappings. It is therefore possible that if the relation between weight and the available symbols is non-arbitrary and can be based on ordinal structure, the preference for sensorimotor communication may be reduced.

To test this prediction, in Experiment 4, we used magnitude-related symbols that bear a natural association with weight and afford a systematic ordinal mapping (small magnitude—light weight; large magnitude—heavy weight). Previous research has shown that there is a general magnitude system in the human brain that processes size-related information from different cognitive and sensorimotor domains (Walsh, 2003). Moreover, it has been proposed that the ability for numerical processing is based on the ability to

perceive size (Henik, Gliksman, Kallai, & Leibovich, 2017). Relatedly, an expectation for larger objects to be heavier has been indirectly demonstrated by the size-weight illusion (Charpentier, 1891; Ernst, 2009; Flanagan & Beltzner, 2000), which shows that the smaller of two equally weighted objects is perceived as heavier.

Based on these previous findings, one can expect that people should easily map numerical symbols to weights since large numbers are naturally associated with larger sizes and, in turn, with heavier weights. Importantly, both numerosity and weight have an ordinal structure, allowing for a systematic mapping between one ordered series to another. In Experiment 4, we made use of this pre-established magnitude-related association between numerosity and weight to test whether co-actors' preference for sensorimotor communication (as observed in the previous experiments) can be shifted to a preference for symbolic communication when the available symbols bear an intrinsic relation to the hidden object property and afford a systematic ordinal mapping.

To this end, we attached numerosity cues (i.e., 1–3 small dots) on the objects in a way that allowed us to distinguish between the previously used modulations of grasp height and modulations that targeted the numerosity cues. As the majority of participants in Experiments 1–3 had grasped heavy objects at the bottom and light objects at the top, we attached the dots in the reverse order such that three dots (that should be associated with “heavy”) were attached at the top and one dot (associated with “light”) was attached at the bottom of the objects (see Fig. 1). If participants in Experiment 4 created mappings between the numerosity cues and the object weights, they would grasp heavy objects at the top and light objects at the bottom. Conversely, if participants disregarded the numerosity cues and continued to use the same grasp height modulations as in the previous experiments, they would grasp heavy objects at the bottom and light objects at the top.

5.1. Method

The method used in Experiment 4 was the same as in the previous experiments, with the following exceptions.

5.1.1. Participants

Ten female and two male volunteers participated in randomly matched pairs (four only-female pairs, $M_{\text{age}} = 22.8$ years, $SD = 2.51$ years, range: 19–26). Two pair members had met before (both were students at the same university).

5.1.2. Apparatus and stimuli

The objects used in Experiment 4 were white as in Experiment 1 (Fig. 1). Red dots (diameter: 1 cm) were attached to each object in the following way: One dot was located at a height of 5 cm, two dots were located at a height of 13.5 cm, and three dots were located at a height of 22 cm. The same arrangement of dots was attached on the two opposite sides of each object such that they could be seen from all angles.

5.1.3. Procedure

Experiment 4 consisted only of a joint condition. An individual baseline was not deemed necessary because the individual baselines of the three previous experiments had all yielded very consistent data, suggesting that people do not grasp objects in a weight-specific manner when acting individually.

5.1.4. Data analysis

In contrast to the previous three experiments, we used the signed instead of the absolute differences in grasp height as our main parameter in Experiment 4, because we now tested a directional hypothesis. We predicted that participants would grasp heavy objects at the top and light objects at the bottom and not vice versa as in the previous experiments. Positive difference values imply that heavy objects are grasped at higher positions than light objects, whereas negative difference values imply the reverse.

We applied the same sampling-without-replacement procedure as used to analyze participants' choice of color in Experiments 2 and 3, except that we now computed the choice of numerosity cue. To this end, we divided the object into three sections, such that a grasp location in a particular section counted as choice of the particular numerosity cue located within this section. Specifically, the object was divided at the heights of 9.75 cm and of 18.25 cm, such that there were 3.75 cm between these division lines and the adjacent numerosity cue, as the dots were attached at 5, 13.5, and 22 cm, respectively, and had a diameter of 1 cm.

As our prediction in Experiment 4 depended on the observation that participants in the previous experiments mostly used a height-weight mapping in which heavy was coded as a low grasp and light as a high grasp, we tested whether in Experiment 4, participants would be significantly less likely to use this mapping direction. Such a finding would indicate that they made use of the numerosity cues instead of the height-weight mapping. Thus, to test whether participants in Experiment 4 used the same or the reverse mapping, we compared the frequencies of participants' preferred mapping direction (i.e., whether they preferred to map a high grasp location to a light or a heavy weight) between Experiment 4 and Experiment 1. The only difference between the objects used in these two experiments was the numerosity cues attached to the objects' surfaces in Experiment 4. Only participants who used height-weight mappings were included in this analysis because the data from participants who used a different system were lacking the relevant direction values. Fisher's exact test was used to determine whether there was a significant difference between participants from Experiment 4 and Experiment 1 in the mapping direction they preferred.

5.2. Results

Due to technical recording errors or procedural errors, 7.9% of trials were excluded from the analysis.

5.2.1. Grasp height

Consistent with previous findings, informed participants adjusted their grasp height to the weight of the grasped object. In contrast to previous findings, most participants grasped heavy objects at higher positions than light objects, as indicated by the positive difference values ($M_{\text{asymmetric}} = 3.13$; $M_{\text{symmetric}} = 2.29$); see Fig. 7a. Participants' grasp height for the medium weight differed significantly from the grasp height for the light weight ($B = -3.33$, $p < .001$) and from the grasp height for the heavy weight ($B = 2.91$, $p = .003$) in the asymmetric condition (see left panel in Fig. 7a). There was also a significant difference between the grasp heights for the medium weight and the light weight ($B = -2.65$, $p = .008$) and between the grasp heights for the medium weight and the heavy weight ($B = 2.08$, $p = .039$) in the symmetric condition (see right panel in Fig. 7a).

As in previous experiments, there was no significant difference between the size of the grasp height differences in the asymmetric and the symmetric condition (medium vs. light: $B = 0.73$, $p = .194$; medium vs. heavy: $B = 0.69$, $p = .220$), indicating that participants consistently used communicative signals independent of their co-actor's knowledge state.

5.2.2. Numerosity choice

The analysis of numerosity choice showed that participants used consistent mappings between different numerosity cues and object weights (see Fig. 7b for descriptive results). In both the asymmetric and the symmetric condition, the three-one-sample t tests for the three weights reached significance, given the Bonferroni-corrected significance level of $p = .008$ accounting for multiple comparisons (asymmetric: light: $t(11) = 4.56$, $p = .001$, Cohen's $d = 1.30$; medium: $t(11) = 6.21$, $p < .001$, Cohen's $d = 1.80$; heavy: $t(11) = 5.09$, $p < .001$, Cohen's $d = 1.45$; symmetric: light: $t(11) = 3.93$, $p = .002$, Cohen's $d = 1.15$; medium: $t(11) = 5.87$, $p < .001$, Cohen's $d = 1.69$; heavy: $t(11) = 3.94$, $p = .002$, Cohen's $d = 1.14$). These results imply that participants consistently placed their grasps onto one specific numerosity section for one specific weight with a frequency above chance level.

5.2.3. Direction of height-weight mapping

In Experiment 4, only two of nine participants preferred to map high grasp locations onto light weights, whereas in Experiment 1, 11 of 11 participants preferred this mapping direction. This difference was statistically significant ($p < .001$, Fisher's exact test). This result suggests that participants in Experiment 4 preferred the reverse mapping direction more than participants in Experiment 1. Whereas in Experiment 1, participants mapped the height of their grasps to the weights of the grasped objects, in Experiment 4 it was not the grasp height that was mapped to weight but the numerosity cues attached at a certain height. Participants grasped light objects at the bottom where one dot was attached, medium objects around the middle (two dots), and heavy objects at the top (three dots); see Fig. 7a.

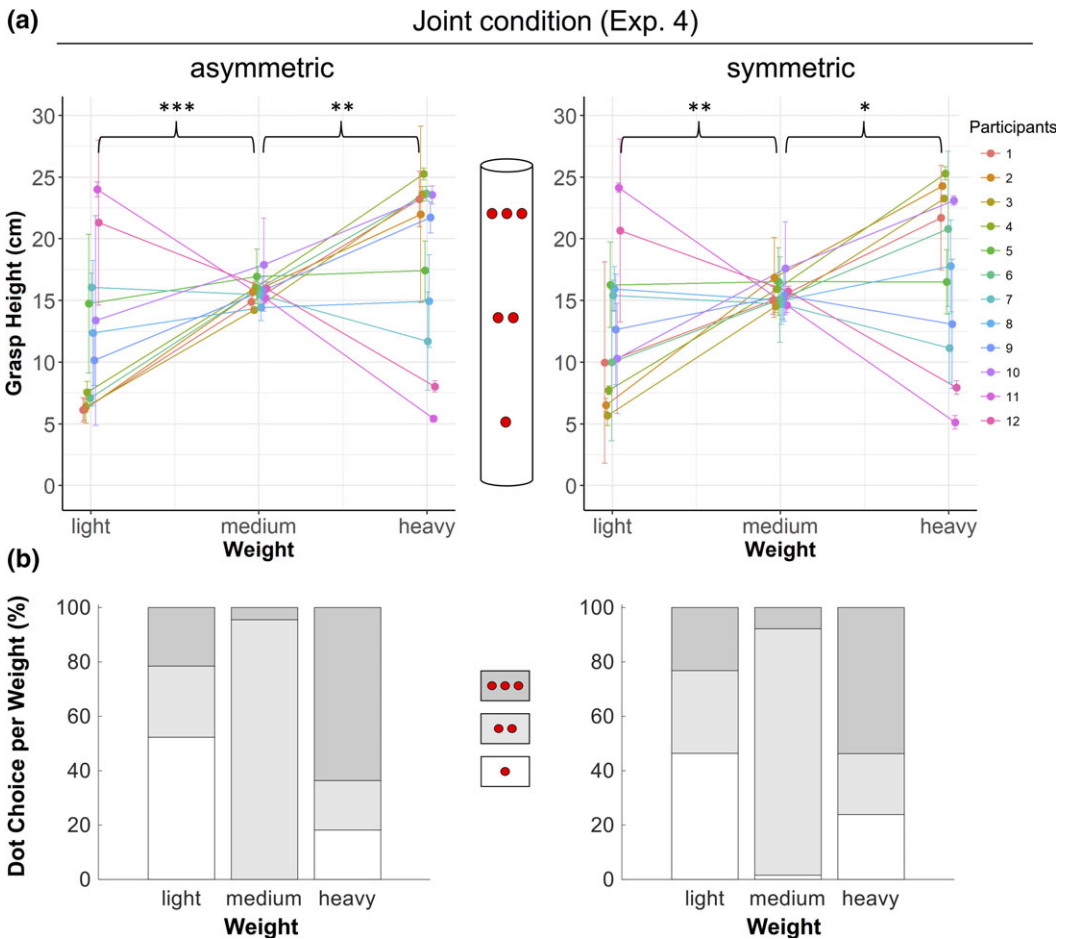


Fig. 7. Mean grasp height in Experiment 4 is shown as a function of object weight and participant. Seven of 12 participants chose grasp positions indicating a numerosity-weight mapping: They grasped light objects at the bottom where one dot was attached, medium objects around the middle (two dots), and heavy objects at the top (three dots). The object centrally depicted in (a) illustrates where the numerosity cues were located relative to participants' grasp positions. Error bars show standard deviations. Panel (b) shows the mean choice of dots (in % of trials) per weight and condition across participants. Dark gray represents three dots, medium gray represents two dots, and white represents one dot. $*p < .05$, $**p < .01$, $***p < .001$.

5.2.4. Matching accuracy

Co-actors in Experiment 4 achieved an accuracy of 81.4%. This value differed significantly from 33% chance performance, $t(5) = 9.21$, $p < .001$, Cohen's $d = 3.76$.

5.2.5. Verbal reports

Of the seven participants who showed the predicted height-weight mapping (i.e., participants 1–4, 6, 9–10; see Fig. 7a), five participants explicitly reported that they mapped the number of dots to the object weight (i.e., 1 dots = light, 2 dots = medium, 3

dots = heavy) by grasping the object at the position where the respective number of dots was attached. They used this mapping strategically to communicate the object weight to their partner. The two other participants reported that they mapped object height to object weight, without mentioning the dots explicitly. Participant 9 only modulated her grasp height in the asymmetric condition but not in the symmetric condition. Participant 5 also reported to have used a height-weight mapping, yet she developed this strategy only very late in the experiment so that her mean grasp height values do not reflect any differences (see Fig. 7a). The two pairs who did not apply the predicted mapping (see participants 11 & 12 and 7 & 8 in Fig. 7a) reported to have used a height-weight mapping (i.e., the same as observed in previous experiments) and to have modulated movement velocity to indicate object weight, respectively.

5.3. Discussion

The results of Experiment 4 provided tentative evidence for our hypothesis that co-actors choose to establish a symbolic communication system by mapping specific symbols to specific object weights if these mappings are not completely arbitrary but rely on natural associations and on matching ordinal structures. In particular, there are intrinsic relations between magnitude concepts such as numerosity, size, and weight (e.g., Charpentier, 1891; Henik et al., 2017) that participants in the current experiment exploited to establish a reliable and consistent communication system. For most of the participants, this numerosity-based symbol system overruled the preference for a height-weight mapping that we had observed in the previous experiments. Participants' verbal reports support these behavioral results as using the dots in a way that associated three dots with heavy, two dots with medium, and one dot with light was the most frequently reported communication strategy.

Future research may investigate whether the numerosity-based communication system in the present task was driven by perceptual common ground or by conceptual common ground (see Clark, 1996). This could be done by testing whether co-actors would also rely on the numerosity cues if their objects had different cue configurations, for example, such that the dots on one actor's object set were attached in the order 3-2-1 from top to bottom, whereas the dots on the co-actor's object set were attached in the order 1-3-2. If the usage of the numerosity system was predominantly driven by perceptual common ground, then co-actors might not rely on numerosity if the numerosity configurations differ perceptually. However, if the usage of the numerosity system was driven by conceptual common ground, then co-actors should again rely on this system, mapping different magnitudes to different weights.

6. General discussion

The aim of this study was to investigate how co-actors involved in a joint action communicate hidden object properties without relying on pre-established communicative

conventions in language or gesture. We hypothesized that they would either use sensorimotor communication by systematically modulating their instrumental movements or that they would communicate in a symbolic manner. Given that most previous research on sensorimotor communication focused on the communication of spatial location, our study addressed the important question of whether the scope of sensorimotor communication can be extended to accommodate the need to create mappings with hidden object properties, and if so, whether this way of communicating is preferred over using symbolic forms of communication. To investigate whether people generally prefer one or the other of these two types of communication systems, or whether situational factors affect their preference, we designed a task that would allow for the emergence of either of the two systems.

The task for two co-actors was to establish a balance on a scale by selecting two objects of equal weight. Only one actor possessed weight information. Her uninformed co-actor needed to choose one of three visually indistinguishable objects that differed in weight, trying to match the weight of her partner's object. The results from Experiments 1–3 consistently showed that informed actors transmitted weight information to their uninformed co-actors by systematically modulating their instrumental actions, grasping objects of different weights at different heights. Although participants in Experiments 2–3 preferred sensorimotor communication even if they had the opportunity to create arbitrary mappings between specific colors and specific weights, Experiment 4 showed that introducing symbols that bear an intrinsic relation to weight and afford mapping based on ordinal structure resulted in a preference switch. The majority of participants now preferred a non-arbitrary, ordinal mapping between numerosity cues and object weights over the previously used grasp height modulations. Across all four experiments, participants managed to create functional communication systems, as indicated by high accuracy values.

Contrary to our predictions that co-actors would only communicate if it was crucial to achieve the joint goal of balancing the scale, participants engaged in communication even if their co-actor did not need the transmitted information. Rather than serving an informative purpose, this persistent use of communication may have served to maintain the overall functionality of the communication system. By consistently using the jointly established communicative signals, communicators might have acknowledged and confirmed their commitment to the joint action (see Michael et al., 2015). Moreover, we observed during data collection that when both co-actors were informed about object weight, the co-actor who acted second often grasped her object at the same height as the co-actor who acted first. Since this copying did not have any informative function, we suggest that participants in the role of the addressee might have adhered to the use of the communicative signals to acknowledge that they understand the meaning of these signals and to confirm that the shared usage of these signals is part of the co-actors' common ground (see Clark & Brennan, 1991, on "grounding" in communication).

The present findings go beyond the typically investigated cases of sensorimotor communication where communicative deviations directly map onto the to-be-communicated spatial properties (e.g., a higher movement amplitude implies a higher grasp location) and

provide initial evidence that sensorimotor communication can accommodate mappings between movement modulations and non-spatial properties which must be created *de novo*. Furthermore, whereas typical cases of sensorimotor communication need not necessarily involve conscious processes (Pezzulo et al., 2013), creating the mappings in the present task required a more strategic, most likely conscious, intention to communicate.

An important question arising from the present findings is whether using differences in grasp height is a communicative strategy specific to weight or whether the same strategy could also be applied to other hidden object properties such as fragility or rigidity. A small survey we conducted post-hoc with a new participant sample ($N = 24$) suggested that the strategy might be weight-specific: When asked where they would spontaneously grasp objects of different weights, the majority of participants replied that they would grasp light objects at the top (18 of 24) and heavy objects at the bottom (19 of 24). When asked about where they would grasp objects of different rigidity (soft vs. hard) or fragility (breakable vs. unbreakable), no similar consistent pattern of responses emerged.

Thus, people's explicit associations seem consistent with the sensorimotor signals used by participants in the present study who most often communicated "light" by grasping objects at the top and "heavy" by grasping objects at the bottom. Together, these findings suggest that grasp height as a signal for weight was not chosen at random (or due to a lack of alternative options) but because people have clear associations between object weight and grasp height. However, the fact that people have these clear associations points to an interesting discrepancy: Although people consistently matched light weight to high grasp positions and heavy weight to low grasp positions when asked explicitly, these associations were not reflected in people's individual grasping behavior, as demonstrated by the individual baseline data obtained in our study.

What is the reason for this mismatch between people's explicit statements and their motor behavior? It is possible that grasp height differences in the individual baselines were absent because the weights of the three objects used in this study did not differ much. Grasp height differences in behavior might only emerge for sufficiently large differences in weight. To address this possibility, we collected additional data for two objects that strongly differed in weight (70 g vs. 1510 g). We asked a new group of individual participants ($N = 6$) to repeatedly grasp these objects in a randomized order. No weight-specific differences in participants' grasp height were found (see Fig. S2 in supplementary material).

However, even though there was no evidence for weight-specific grasping differences in our data, it is still likely that people's explicit associations are based on certain weight-specific action affordances. For example, people have a tendency to grasp very heavy (and possibly slippery) objects—such as a heavy cardboard box full of books—by placing their hands underneath the object to make sure it does not slip out of their hands. No such tendency exists for light objects, as the danger of dropping them is much smaller. Thus, objects of heavy weight and potential slipperiness might form the anchor for people's explicit associations between low grasps and heavy weights, and between high grasps and light weights, as derived by contrast. These associations are then translated into the sensorimotor medium for communicative purposes. The reason we did not

observe any weight-specific differences in participants' individual grasping behavior could be that the objects were simply not heavy (and slippery) enough. The grasping tendencies described above might only emerge for specific objects that have the right combination of heaviness and slipperiness, for example, heavy cardboard boxes.

What can serve as a basis for bootstrapping novel communication systems in the absence of pre-established communicative conventions? In the case of sensorimotor communication, an alternative to relying on action affordances could be relying on naturally occurring regularities in the physical and cultural environment. People exhibit consistent associations between stimuli features from different sensory modalities; for example, they consistently associate high-pitched sounds with small, bright, and sharp-edged objects (e.g., Hubbard, 1996; Marks, 1987, 1989; see also Deroy & Spence, 2016). These “crossmodal correspondences” (for a comprehensive review, see Spence, 2011) are based on natural statistical regularities (e.g., between the size of an object and its resonance frequency; Bee, Perrill, & Owen, 2000; Coward & Stevens, 2004), structural associations (e.g., between magnitude-related stimuli features such as size and loudness; Smith & Sera, 1992; Walsh, 2003), and semantically mediated associations (e.g., between auditory pitch and visual elevation; Martino & Marks, 1999; also see Walker & Walker, 2012). Interestingly, the association between higher pitches and higher positions in space has been shown to exist even in pre-linguistic infants, indicating that language builds on preexisting mappings (Dolscheid, Hunnius, Casasanto, & Majid, 2014; Walker, Bremner, et al., 2010).

Relatedly, work on “mental metaphors” (Casasanto, 2009) suggests that even abstract ideas often have a basis in how people experience their physical and cultural environment (see Lakoff & Johnson, 2008). Supporting this view, studies have demonstrated that people's perceptuomotor experience affects their mental representations, for instance such that right-handers associate rightward space with *good* and leftward space with *bad* because they can act more fluently with their dominant hand in the right side of space (Casasanto, 2009; see also Casasanto & Dijkstra, 2010). In the context of our study, people could have drawn on the basic metaphor of More is Up and Less is Down (e.g., Lakoff, 2012), developing a communicative mapping between high grasps and heavy weights (and conversely between low grasps and light weights)—a mapping we did observe, albeit less frequently than its reverse.

Based on the research outlined above, one may hypothesize that naturally occurring associations may serve as “natural conventions” (Deroy & Spence, in press) when creating novel communication systems, such that co-actors could build on crossmodal correspondences or mental metaphors to create communicative mappings. In line with this idea, it has been suggested that mappings between naturally associated dimensions are more easily identified than mappings between unrelated dimensions (Coward & Stevens, 2004).

Building on the present findings, it would be interesting to explore how people create novel communication systems to transmit information about other hidden object properties, for example fragility or rigidity, and what serves as the basis for these systems. Future research may also look at properties that are not quantity-based (e.g., using liquid, solid, or particulate contents), as it is possible that communicating these properties generally differs from communicating quantity-based properties such as weight. We used

weight in our study to provide evidence in one domain that sensorimotor communication can be used to communicate hidden object properties. More research is needed to show whether our findings can be generalized to other hidden object properties.

Why did participants in the present study modulate the end state of their action rather than the movement phase, as observed in previous studies (e.g., Sacheli et al., 2013; Vesper et al., 2016)? The preference for grasp height as a communicative signal may be explained by its static (instead of dynamic) nature which allows observers more time to recognize the signal compared to modulations in movement velocity or maximal movement amplitude (e.g., Vesper & Richardson, 2014) which are dynamic in nature and fade quickly. It is also possible that people preferred modulations of the action end state over modulations of the movement phase because these were more clearly recognizable as a communicative signal (see Vesper, Schmitz, & Knoblich, 2017). Note, however, that in the present study, differences in participants' grasp height were highly correlated with differences in the peak height of participants' reach-to-grasp trajectories. Thus, observers could have already picked up on the communicative signal during the approach phase and then confirmed their first impression by observing their partner's final grasp position. Based on participants' verbal reports, it can be safely assumed that it was the final grasp height that participants aimed to make distinctive for communication, yet while doing so they naturally also modulated their reach-to-grasp trajectories.

Using differences in grasp height as a communicative signal evokes the prediction that the larger the differences displayed by informed actors, the better uninformed co-actors should be able to recognize the signal. This should be evident in the form of a high correlation between the size of the average difference in grasp height and co-actors' matching accuracy. However, we could not empirically confirm this prediction because the correlations computed from the present data were uninformative due to the very small variability in accuracy between pairs ($M_{\text{Exp1-4}} = 89\%$, $SD_{\text{Exp1-4}} = 14\%$).

It is an open question whether the grasp-based communication system participants established in this study is generalizable beyond the dyadic interaction in which it was created. This question relates to previous work showing that people adjust to their conversational partners' linguistic preferences by aligning on grammatical structure, task-relevant vocabulary, and referring expressions (e.g., Brennan & Clark, 1996; Brown-Schmidt, 2009; Fusaroli et al., 2012; Garrod & Pickering, 2004; Mills, 2011). Moreover, people adapt their own communicative behavior (i.e., speech and gesture) to their conversational partners' knowledge and beliefs ("audience design"; e.g., Clark, 1996; Clark & Murphy, 1982; Holler & Stevens, 2007; Horton & Keysar, 1996; Newman-Norlund et al., 2009; Özyürek, 2002). A recent study (Vesper, Schmitz, et al., 2017) examined partner-specificity in sensorimotor communication, showing that a sensorimotor communication system created by participants in an interactive task was stable and generalizable to an offline setting rather than tailored to a specific interaction partner. This provides a hint that the communication system established in the present task might be generalizable as well.

Previous work on partner-specificity in conversation is consistent with our finding that almost all task partners in this study aligned on the same communicative strategy. In principle, it is possible for co-actors to use different communicative signals, as long as

they understand the meaning of each other's signals. However, there were only two (of 24) pairs where task partners did not use a common strategy. Using a common set of communicative signals may have facilitated coordination, as suggested by previous research on conversation showing that task partners who used the same (task-relevant) language achieved higher levels of coordination (Fusaroli et al., 2012).

Furthermore, it is interesting that most co-actors established a functional communication system very rapidly, needing only a few trials to develop and align on a certain set of signals (for similar findings, see Vesper, Schmitz, et al., 2017). Notably, symmetric trials where co-actors could rely on common ground (Clark, 1996) might have played a critical role in this process: By using communicative signals when there was mutual knowledge about the weight of a given object, co-actors could make manifest the intended meaning of a specific modulation and thereby reassure each other that they were on the same page. This way, they could rapidly build up new common ground and establish a functional communication system.

Regarding our initial question of which type of communication system co-actors preferentially establish to communicate hidden object properties, the present findings suggest that their preference depends on situational factors. Specifically, co-actors showed a consistent preference for sensorimotor communication when the alternative option was to create arbitrary mappings between specific colors and specific object weights. Yet when the mappings were not arbitrary but based on an intrinsic relation and on corresponding ordinal structures between sign and referent, co-actors' preference switched to the use of these mappings. In the present study, we relied on the intrinsic magnitude-based relation between weight and numerosity using 1–3 dots as numerosity cues. However, the magnitude dimension could be represented in many other ways, from using numerals (e.g., “1,” “2,” “3”) to gradients of color (e.g., a gradient from white over gray to black, see Walker, Francis, et al., 2010). It is likely that the present findings would generalize to other instantiations of magnitude, potentially even to those that do not bear any intrinsic relation to weight—such as a series of smiley faces with different degrees of happiness—as long as they afford a systematic mapping based on ordinal structure. Future studies should determine whether the observed preference for numerosity over color can be ascribed to the more motivated, less arbitrary relation between numerosity and weight, or solely to the fact that numerosity and weight both contain an ordinal structure.

Whether mapping specific colors or numerosity cues to specific weights can be described as a more *symbolic* form of communication compared to the grasp height system can be regarded as a question of definition. Arguably, both types of communication systems can be seen as symbolic in the basic sense that certain elements stand for, or signify, certain other elements. Going beyond this very broad definition, symbols are often contrasted with icons (e.g., Allwood, 2002): Whereas icons and their referents are related by similarity, the relation between symbols and their referents is arbitrary; for example, it can be determined by convention (also see the distinction between icons, indices, and symbols introduced by Peirce; for example, the Peirce Edition Project, 1984).

Given this distinction between symbols and icons, the mapping between specific numerosity cues and specific weights can be regarded as iconic, as it relies on a shared

quality between sign and referent (i.e., numerosity and weight share an ordinal magnitude dimension). Research in experimental semiotics suggests that icons can evolve into symbols through an interactive grounding process between producer and receiver of the signs (Garrod et al., 2007). In the beginning of a communicative interaction, the informational content of an iconic sign is grounded in the sign's physical appearance but after repeated interactive usage, the icon evolves into a symbol (i.e., an often simpler and more abstract representation) with its informational content being grounded in the users' shared history (Galantucci, 2009; Garrod et al., 2007).

In contrast to the iconic numerosity signs, the colors used in our study do not bear any similarity with weight. Thus, the mapping between specific colors and specific weights can be regarded as arbitrary, in this respect resembling typical symbolic word-meaning mappings in human language. Finally, the relation between specific grasp heights and specific weights is not one of similarity, yet it is not completely arbitrary as it is based on people's prior associations, which are possibly grounded in weight-specific action affordances (see discussion above).

Thus, the three different signs (i.e., numerosity, color, grasp height) connect to their referent (weight) in different ways (e.g., see Peirce; the Peirce Edition Project, 1998). Whereas there is a preexisting (magnitude-related) similarity between numerosity and weight, there is no prior connection between color and weight. This qualifies the former as iconic and the latter as symbolic relation. The connection between grasp height and weight seems neither of an iconic nor of a symbolic nature, as it is not based on similarity yet it is also not arbitrary but built on preexisting associations. The present study has shown that co-actors prefer the latter type of mappings over arbitrary (color-weight) mappings but not over iconic (numerosity-weight) mappings. Together, these findings not only extend previous research on sensorimotor communication, but they also provide novel insights into how communication emerges out of the need to coordinate.

Acknowledgments

This work was supported by the European Research Council under the European Union's Seventh Framework Program (FP7/2007-2013)/ERC grant agreement no. 609819, SOMICS, and by ERC grant agreement no. 616072, JAXPERTISE. We thank Dávid Csűrös, Fruzsina Elekes, and Anna Mária Lisincki for their help with data collection and stimuli creation, as well as for their helpful comments. We also thank Riccardo Fusaroli, Kensy Cooperrider, and three anonymous reviewers for their helpful comments and suggestions.

Notes

1. In the pilot study ($N = 8$ pairs), the majority of participants used grasp height (i.e., high vs. low grasp) to disambiguate light versus heavy weight. The second most preferred communicative signal was the modulation of grasp type (precision vs.

power grip), and another strategy was the modulation of movement velocity. Note that the grasp type strategy was only suitable for signaling one of two alternatives (e.g., heavy or light). Thus, this strategy would not have worked in our main study because three weights needed to be distinguished.

2. Prior literature suggests that people exhibit intuitive brightness-weight mappings, automatically associating darker shades with heavier objects and brighter shades with lighter objects (Walker, Francis, & Walker, 2010).
3. Note that this observation is specific to the three colors used in our study. More generally, ordinal structure can be found in certain color sets, for example, in the spectrum of rainbow colors (ROYGBV) or in a set of colors ranging from darker to lighter shade.

References

- Aglioti, S. M., Cesari, P., Romani, M., & Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nature Neuroscience*, *11*(9), 1109–1116.
- Alaerts, K., Swinnen, S. P., & Wenderoth, N. (2010). Observing how others lift light and heavy objects: Which visual cues mediate the encoding of muscular force in the primary motor cortex? *Neuropsychologia*, *48*, 2082–2090.
- Allwood, J. (2002). Bodily communication dimensions of expression and conflict. In B. Granström, D. House, & I. Karlsson (Eds.), *Multimodality in language and speech systems* (pp. 7–26). Dordrecht, the Netherlands: Springer.
- Bahrami, B., Olsen, K., Latham, P. E., Roepstorff, A., Rees, G., & Frith, C. D. (2010). Optimally interacting minds. *Science*, *329*(5995), 1081–1085.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Becchio, C., Sartori, L., Bulgheroni, M., & Castiello, U. (2008). The case of Dr. Jekyll and Mr. Hyde: A kinematic study on social intention. *Consciousness and Cognition*, *17*, 557–564.
- Bee, M. A., Perrill, S. A., & Owen, P. C. (2000). Male green frogs lower the pitch of acoustic signals in defense of territories: A possible dishonest signal of size? *Behavioral Ecology*, *11*(2), 169–177.
- Bingham, G. P. (1987). Kinematic form and scaling: Further investigations on the visual perception of lifted weight. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(2), 155–177.
- Bosbach, S., Cole, J., Prinz, W., & Knoblich, G. (2005). Inferring another's expectation from action: The role of peripheral sensation. *Nature Neuroscience*, *8*(10), 1295–1297.
- Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1482–1493.
- Brown-Schmidt, S. (2009). Partner-specific interpretation of maintained referential precedents during interactive dialog. *Journal of Memory and Language*, *61*, 171–190.
- Buckingham, G., Cant, J. S., & Goodale, M. A. (2009). Living in a material world: How visual cues to material properties affect the way that we lift objects and perceive their weight. *Journal of Neurophysiology*, *102*(6), 3111–3118.
- Buckingham, G., & Goodale, M. A. (2010). Lifting without seeing: The role of vision in perceiving and acting upon the size weight illusion. *PLoS ONE*, *5*(3), e9709.
- Candidi, M., Curioni, A., Donnarumma, F., Sachelì, L. M., & Pezzulo, G. (2015). Interactional leader-follower sensorimotor communication strategies during repetitive joint actions. *Journal of the Royal Society Interface*, *12*, 20150644.

- Casasanto, D. (2009). Embodiment of abstract concepts: Good and bad in right-and left-handers. *Journal of Experimental Psychology: General*, *138*(3), 351–367.
- Casasanto, D., & Dijkstra, K. (2010). Motor action and emotional memory. *Cognition*, *115*(1), 179–185.
- Casile, A., & Giese, M. A. (2006). Nonvisual motor training influences biological motion perception. *Current Biology*, *16*(1), 69–74.
- Cavallo, A., Koul, A., Ansuini, C., Capozzi, F., & Becchio, C. (2016). Decoding intentions from movement kinematics. *Scientific Reports*, *6*, 37036.
- Charpentier, A. (1891). Analyse experimentale de quelques elements de la sensation de poids. *Archive de Physiologie Normales Pathologiques*, *3*, 122–135.
- Clark, H. H. (1996). *Using language*. Cambridge, UK: Cambridge University Press.
- Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127–149). Washington, DC: American Psychological Association.
- Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, *50*(1), 62–81.
- Clark, H. H., & Murphy, G. L. (1982). Audience design in meaning and reference. *Advances in Psychology*, *9*, 287–299.
- Clark, H. H., & Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science*, *13*(2), 259–294.
- Coward, S. W., & Stevens, C. J. (2004). Extracting meaning from sound: Nomic mappings, everyday listening, and perceiving object size from frequency. *The Psychological Record*, *54*(3), 349.
- Cross, E. S., Hamilton, A. F. D. C., & Grafton, S. T. (2006). Building a motor simulation de novo: Observation of dance by dancers. *NeuroImage*, *31*(3), 1257–1267.
- de Ruiter, J. P., Noordzij, M., Newman-Norlund, S., Hagoort, P., & Toni, I. (2007). On the origin of intentions. *Attention & Performance*, *XXII*, 593–610.
- de Saussure, F. (1959). *Course in general linguistics*. C. Bally & A. Sechehaye with A. Riedlinger (Eds.) & W. Baskin (Trans.). New York: Philosophical Library.
- Deroy, O., & Spence, C. (2016). Crossmodal correspondences: Four challenges. *Multisensory Research*, *29* (1–3), 29–48.
- Deroy, O., & Spence, C. (in press). Crossing the senses: How crossmodal correspondences shape perception and culture. Oxford: Oxford University Press.
- Dolscheid, S., Hunnius, S., Casasanto, D., & Majid, A. (2014). Prelinguistic infants are sensitive to space-thing associations found across cultures. *Psychological Science*, *25*(6), 1256–1261.
- Duff, M. C., Hengst, J., Tranel, D., & Cohen, N. J. (2006). Development of shared information in communication despite hippocampal amnesia. *Nature Neuroscience*, *9*(1), 140–146.
- Ernst, M. O. (2009). Perceptual learning: Inverting the size–weight illusion. *Current Biology*, *19*(1), R23–R25.
- Flanagan, J. R., & Beltzner, M. A. (2000). Independence of perceptual and sensorimotor predictions in the size–weight illusion. *Nature Neuroscience*, *3*(7), 737–741.
- Fusaroli, R., Bahrami, B., Olsen, K., Roepstorff, A., Rees, G., Frith, C., & Tylén, K. (2012). Coming to terms: Quantifying the benefits of linguistic coordination. *Psychological Science*, *23*(8), 931–939.
- Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive Science*, *29*(5), 737–767.
- Galantucci, B. (2009). Experimental semiotics: A new approach for studying communication as a form of joint action. *Topics in Cognitive Science*, *1*(2), 393–410.
- Galantucci, B., Garrod, S., & Roberts, G. (2012). Experimental semiotics. *Language and Linguistics Compass*, *6*, 477–493.
- Garrod, S., Fay, N., Lee, J., Oberlander, J., & MacLeod, T. (2007). Foundations of representation: Where might graphical symbol systems come from? *Cognitive Science*, *31*(6), 961–987.
- Garrod, S., Fay, N., Rogers, S., Walker, B., & Swoboda, N. (2010). Can iterated learning explain the emergence of graphical symbols? *Interaction Studies*, *11*(1), 33–50.

- Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, 8, 8–11.
- Gordon, A. M., Forssberg, H., Johansson, R. S., & Westling, G. (1991). Visual size cues in the programming of manipulative forces during precision grip. *Experimental Brain Research*, 83(3), 477–482.
- Grèzes, J., Frith, C., & Passingham, R. E. (2004a). Inferring false beliefs from the actions of oneself and others: An fMRI study. *NeuroImage*, 21, 744–750.
- Grèzes, J., Frith, C., & Passingham, R. E. (2004b). Brain mechanisms for inferring deceit in the actions of others. *The Journal of Neuroscience*, 24(24), 5500–5505.
- Hamilton, A., Joyce, D. W., Flanagan, J. R., Frith, C. D., & Wolpert, D. M. (2007). Kinematic cues in perceptual weight judgment and their origins in box lifting. *Psychological Research*, 71, 13–21.
- Healey, P. G., Swoboda, N., Umata, I., & Katagiri, Y. (2002). Graphical representation in graphical dialogue. *International Journal of Human-Computer Studies*, 57(4), 375–395.
- Healey, P. G., Swoboda, N., Umata, I., & King, J. (2007). Graphical language games: Interactional constraints on representational form. *Cognitive Science*, 31(2), 285–309.
- Henik, A., Gliksman, Y., Kallai, A., & Leibovich, T. (2017). Size perception and the foundation of numerical processing. *Current Directions in Psychological Science*, 26(1), 45–51.
- Holler, J., & Stevens, R. (2007). The effect of common ground on how speakers use gesture and speech to represent size information. *Journal of Language and Social Psychology*, 26(1), 4–27.
- Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, 59(1), 91–117.
- Hubbard, T. L. (1996). Synesthesia-like mappings of lightness, pitch, and melodic interval. *The American Journal of Psychology*, 109(2), 219–238.
- Knoblich, G., & Jordan, J. S. (2003). Action coordination in groups and individuals: Learning anticipatory control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(5), 1006–1016.
- Lakoff, G. (2012). Explaining embodied cognition results. *Topics in Cognitive Science*, 4(4), 773–785.
- Lakoff, G., & Johnson, M. (2008). *Metaphors we live by*. Chicago: University of Chicago Press.
- Manera, V., Becchio, C., Cavallo, A., Sartori, L., & Castiello, U. (2011). Cooperation or competition? Discriminating between social intentions by observing prehensile movements. *Experimental Brain Research*, 211, 547–556.
- Marks, L. E. (1987). On cross-modal similarity: Auditory-visual interactions in speeded discrimination. *Journal of Experimental Psychology: Human Perception and Performance*, 13(3), 384–394.
- Marks, L. E. (1989). On cross-modal similarity: The perceptual structure of pitch, loudness, and brightness. *Journal of Experimental Psychology: Human Perception and Performance*, 15(3), 586–602.
- Martino, G., & Marks, L. E. (1999). Perceptual and linguistic interactions in speeded classification: Tests of the semantic coding hypothesis. *Perception*, 28(7), 903–923.
- MATLAB. (2014a). The MathWorks, Inc., Natick, MA, United States.
- MATLAB. (2013b). The MathWorks, Inc., Natick, MA, United States.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.
- Michael, J., Sebanz, N., & Knoblich, G. (2015). The sense of commitment: A minimal approach. *Frontiers in Psychology*, 6, 1968.
- Michael, J., Sebanz, N., & Knoblich, G. (2016). Observing joint action: Coordination creates commitment. *Cognition*, 157, 106–113.
- Mills, G. (2011). The emergence of procedural conventions in dialogue. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 471–476). Austin, TX: Cognitive Science Society.
- Misyak, J., Noguchi, T., & Chater, N. (2016). Instantaneous conventions: The emergence of flexible communicative signals. *Psychological Science*, 27, 1550–1561.
- Newman-Norlund, S. E., Noordzij, M. L., Newman-Norlund, R. D., Volman, I. A., De Ruiter, J. P., Hagoort, P., & Toni, I. (2009). Recipient design in tacit communication. *Cognition*, 111(1), 46–54.

- Özyürek, A. (2002). Do speakers design their cospeech gestures for their addressees? The effects of addressee location on representational gestures. *Journal of Memory and Language*, 46(4), 688–704.
- Pasupathi, M., Carstensen, L. L., Levenson, R. W., & Gottman, J. M. (1999). Responsive listening in long-married couples: A psycholinguistic perspective. *Journal of Nonverbal Behavior*, 23(2), 173–193.
- Pezzulo, G., & Dindo, H. (2011). What should I do next? Using shared representations to solve interaction problems. *Experimental Brain Research*, 211(3–4), 613–630.
- Pezzulo, G., Donnarumma, F., & Dindo, H. (2013). Human sensorimotor communication: A theory of signaling in online social interactions. *PLoS ONE*, 8, e79876.
- Project, P. E. (Ed.) (1984). *The writings of Charles S. Peirce: A chronological edition*, Vol. 2. Bloomington, IN: Indiana University Press.
- Project, P. E. (Ed.) (1998). *The essential peirce*, Vol. 2. Bloomington, IN: Indiana University Press.
- R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: <https://www.R-project.org/>.
- Runeson, S., & Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General*, 112(4), 585.
- Sacheli, L., Tidoni, E., Pavone, E., Aglioti, S., & Candidi, M. (2013). Kinematics fingerprints of leader and follower role-taking during cooperative joint actions. *Experimental Brain Research*, 226, 473–486.
- Sartori, L., Becchio, C., & Castiello, U. (2011). Cues to intention: The role of movement information. *Cognition*, 119, 242–252.
- Schegloff, E. A. (1968). Sequencing in conversational openings. *American Anthropologist*, 70(6), 1075–1095.
- Scott-Phillips, T. C., Kirby, S., & Ritchie, G. R. (2009). Signalling signalhood and the emergence of communication. *Cognition*, 113, 226–233.
- Smith, L. B., & Sera, M. D. (1992). A developmental analysis of the polar structure of dimensions. *Cognitive Psychology*, 24(1), 99–142.
- Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, 73(4), 971–995.
- Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition*. Oxford, UK: Blackwell.
- Tylén, K., Weed, E., Wallentin, M., Roepstorff, A., & Frith, C. D. (2010). Language as a tool for interacting minds. *Mind & Language*, 25(1), 3–29.
- Urgesi, C., Moro, V., Candidi, M., & Aglioti, S. M. (2006). Mapping implied body actions in the human motor system. *Journal of Neuroscience*, 26(30), 7942–7949.
- Vesper, C., Abramova, E., Bütepage, J., Ciardo, F., Crossey, B., Effenberg, A., Hristova, D., Karlinsky, A., McEllin, L., Nijssen, S. R. R., Schmitz, L., & Wahn, B. (2017). Joint action: Mental representations, shared information and general mechanisms for coordinating with others. *Frontiers in Psychology*, 7, 2039.
- Vesper, C., & Richardson, M. (2014). Strategic communication and behavioral coupling in asymmetric joint action. *Experimental Brain Research*, 232, 2945–2956.
- Vesper, C., Schmitz, L., & Knoblich, G. (2017). Modulating action duration to establish non-conventional communication. *Journal of Experimental Psychology: General*, 146(12), 1722–1737.
- Vesper, C., Schmitz, L., Safra, L., Sebanz, N., & Knoblich, G. (2016). The role of shared visual information for joint action coordination. *Cognition*, 153, 118–123.
- Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. P. (2010). Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. *Psychological Science*, 21(1), 21–25.
- Walker, P., Francis, B. J., & Walker, L. (2010). The brightness-weight illusion. *Experimental Psychology*, 57, 462–469.
- Walker, P., & Walker, L. (2012). Size–brightness correspondence: Crosstalk and congruity among dimensions of connotative meaning. *Attention, Perception, & Psychophysics*, 74(6), 1226–1240.
- Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, 7(11), 483–488.

- Wilson, D., & Sperber, D. (2004). Relevance theory. In L. R. Horn & G. Ward (Eds.), *The handbook of pragmatics* (pp. 607–632). Oxford, UK: Blackwell.
- Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, *131*(3), 460–473.
- Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1431), 593–602.

Supporting Information

Additional Supporting Information may be found online in the supporting information section at the end of the article:

Fig. S1. Illustration of the sampling-without-replacement procedure for the color choice analysis.

Fig. S2. Data from the additional control experiment.

Data S1. Trial-by-trial accuracy.