University of Birmingham
Research at Birmingham

Decoding gestural iconicity

Hassemer, Julius; Winter, Bodo

DOI:
10.1111/cogs.12680

License:
Other (please specify with Rights Statement)

Document Version
Peer reviewed version

Citation for published version (Harvard):
Hassemer, J & Winter, B 2018, 'Decoding gestural iconicity', *Cognitive Science*, vol. 42, no. 8, pp. 3034-3049.
https://doi.org/10.1111/cogs.12680

Link to publication on Research at Birmingham portal

Publisher Rights Statement:
This is the peer reviewed version of the following article: [FULL CITE], which has been published in final form at [Link to final article using the DOI]. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

•Users may freely distribute the URL that is used to identify this publication.
•Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
•User may use extracts from the document in line with the concept of ‘fair dealing’ under the Copyright, Designs and Patents Act 1988 (?)
•Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 20. Apr. 2024

Decoding gestural iconicity

Julius Hassemer[1,2] & Bodo Winter[3]

[1] Universidade de São Paulo,
Faculdade de Filosofia, Letras e Ciências Humanas

[2] Humboldt-Universität zu Berlin,
Berlin School of Mind and Brain

[3] University of Birmingham,
Department of English Language and Applied Linguistics

Abstract

Representational gestures are used ubiquitously to depict ideas in an iconic fashion, such as when holding up the thumb and the index finger at a certain distance to indicate the size of a matchstick. However, the process by which a physical hand configuration is mentally transformed into abstract spatial information is not well understood. We present a series of experiments that investigate how people decode the physical form of an articulator to derive imaginary geometrical constructs, which are embraced in our use of “gesture form”. We provide quantitative evidence for several key properties that play a role in this process. First, “profiling”, the ability to focus on a structural sub-unit within the complex form of the gesturing hand. Second, “perspective”, for which we show that one and the same handshape seen from different perspectives can lead to different spatial interpretations. Third, “selectivity”, the fact that gestures focus on certain spatial features at the expense of others. Our results provide a first step toward mapping out the process of how representational gestures make the communication of spatial information possible.

Keywords: gesture; selective depiction; profiling; perspective; visual perception

1. Introduction

When people talk, they gesture. In social interactions, gestures serve core communicative functions (Kendon, 2004; Kok, Bergmann, Cienki, & Kopp, 2016), including the depiction of concrete and abstract concepts (Cienki, 1998; Müller, 1998; Goldin-Meadow, 2003) and the facilitation of lexical access (Krauss, Morrel-Samuels, & Colasante, 1991). One of the most common gesture categories are so-called iconic or representational gestures (McNeill, 1992; Kendon, 2014). Iconicity refers to signs that resemble what they express, as when showing the extent of an object by holding two hands at a specific distance to one another. Gestural iconicity has been implicated in a number of important findings in cognitive science. For example, representational gestures have the capacity to change people's temporal concepts (Jamalian & Tversky, 2012; Lewis & Stickles, 2017), and they facilitate the learning of mathematical concepts (Goldin-Meadow, 2003; Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001; Goldin-Meadow, Cook, & Mitchell, 2009).

Despite representational gestures being a major topic in gesture research (McNeill, 1992, 2005; Goldin-Meadow, 2003; Müller, 2004; Kendon, 2004; Streeck, 2008; Mittelberg, 2014), little is known about the cognitive processes by which people decode iconic information from these gestures. Iconicity is not based on a mapping "between objectively defined forms and objectively determined scenes" (Wilcox, 2004: 123; see also Emmorey, 2014: 1). Instead, iconicity in gesture is based on a mental representation that is derived from the physical articulators. Hassemer (2016) distinguishes between "physical form", the actual physical characteristics of the articulator (e.g., hand shape, movement), and "gesture form" including and foregrounding imaginary spatial features such as the mental image of lines (1D) as part of many pointing gestures or tactile surfaces (2D) presented as being in contact with an imaginary object for handling gestures. In this paper we will follow the approach common to Talmy (2018: Ch. 5) and Hassemer (2016; Hassemer & McCleary, under rev.), who describe gesture semantics in terms of schematic spatial structures that are used systematically across different gestures types.

Consider, for example, how Winter, Perlman and Matlock (2014: 388) described a so-called "precision grip" gesture for a small numerical quantity (index finger and thumb close to each other) as a gesture that resembles "holding a small pellet". This verbal paraphrase neglects a whole swath of cognitive steps that must happen in order to decode the meaning of the gesture. How, for example, does the onlooker actually know what fingers to focus on? How is the spatial construct of size derived from the physical configuration of index finger and thumb, which at a bare minimum involves recognizing the space between the index finger and thumb as relevant in this gesture (see also Mittelberg & Waugh, 2009)? The description also ignores the selectivity of iconicity: The onlooker somehow understands the gesture to be "about" size, and not about other characteristics, such as shape. Iconic gestures only provide stripped down, schematic, and highly focused representations of the ideas they depict (Kita, Alibali, & Chu, 2017). The process by which these mental representations are constructed is little understood.

This paper presents a series of experiments that provides a window into the cognitive processes by which gesture form is decoded from physical form. As a model system, we exploit a group of gestures variously discussed as "precision grip" or "ring" gestures. These gestures are used to express "perfection, correctness or exactness" (Kendon, 2004: 227, 238; De Jorio, 2001: 321-322; Lempert, 2011), but also a precise small number (Winter et al., 2014), or "zero" "worthlessness" (Calbris, 2011: 19-21). Important differences in gesture form are shown in Figure 1.

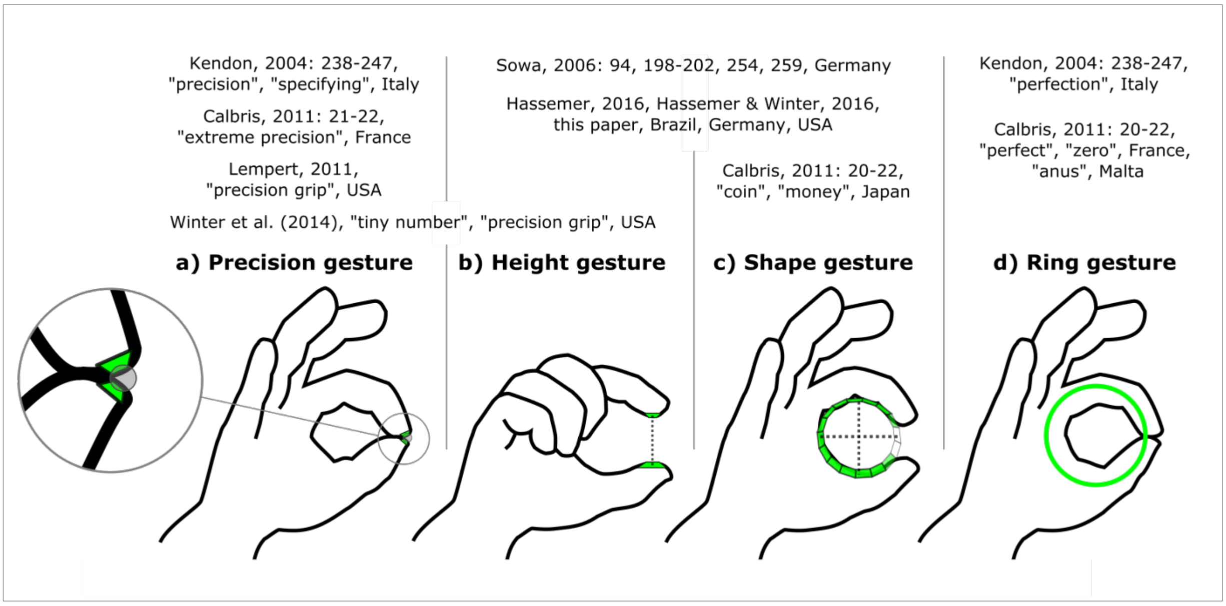

Figure 1. Some gestures that profile the index finger and thumb in different ways with exemplary handshapes; the highlights indicate the profiled areas; a) two very small surfaces at the finger tips, b) two small surfaces at the finger pads, c) one continuous surface at the tactile inside of the fingers, d) one linear circular profile along the fingers' axes

The differences between the gestures in Figure 1 can be characterized in terms of "profiling", a cognitive operation in which attention is directed to a sub-part of a communicative structure (Langacker, 2008: 66). More specifically, "profiling" as part of gesture form analysis (Hassemer, 2016; Hassemer & McCleary under rev.) describes the two initial steps of gesture conceptualization. The first cognitive operation, articulator profiling, focuses on one articulator unit (one or multiple articulators; Hassemer, 2009) within all the potential expressive body parts. In Figure 1a these highlighted body portions are the finger tips; in 1b they are the fingers' distal phalanges; in 1c-d the whole thumb and index finger are highlighted. The second operation, shape profiling, abstracts away from the physical articular and thus more narrowly focuses on one shape feature of the profiled articulator. For the gestures used in this paper's experiments, the shape profile consists of two-dimensional surfaces, which are central in the concept of holding something. For

Figure 1a-c the tactile surfaces are in contact with the imaginary object (see also Sowa, 2006). In Figure 1d there is no contact to an imaginary object, but the shape represents a one-dimensional line coming full circle (see Talmy, 2000: Ch. 1).

Importantly, the imaginary size or shape features of gesture form depend on profiling, which in turn means that selective depiction (the underspecification of certain features) also depends on profiling. Moreover, in this particular case, the meaning of the gesture depends on the visibility of the shape profile, which is affected by the non-profiled articulators. If, for example, the middle, ring and little finger (fingers 3-5) are curled in, then the C-shaped surface within the index finger and thumb (fingers 1-2) is obstructed and not profiled anymore, resulting in a size rather than shape-focused interpretation of the gesture (compare Figure 1b to 1c). This contrast in the physical position of fingers 3-5 will be the core manipulation in our experiments. Experiment 1 provides a conceptual replication-extension of Hassemer and Winter (2016), providing a new test case and highlighting the critical role of imaginary forms, whose effect outplays that of the forms that are physically seen. In experiment 2, we investigate the effect of visual perspective, since the prominence of the shape profile is predicted to shift when it appears only in the background or is occluded by other fingers (Experiment 2). Finally, we demonstrate how profiling leads to selectivity in iconicity (Experiment 3).

2. Experiments 1a and 1b: Production experiments

2.1. Methods

Hassemer and Winter (2016) showed participants a Styrofoam sphere of six centimeters in diameter. Participants were asked to indicate the height and then the shape of the object (in randomized order) using the index finger and thumb of one hand. We found that when asked the “height” question, the majority of participants curled fingers 3-5 in. Here, we follow up on this study, providing an even stronger test for the role of profiling. We reasoned that if participants were holding a small distractor object in their hand, they would be more likely to move it out of the gesturing hand for shape rather than height gestures because the distractor object, as well as fingers 3-5 holding it, would be seen as obstructing the C-shape presentation. We placed two "distractor dice" (approx. 3 centimeters side length) into each of the participant's hands and asked participants to throw both dice in the air simultaneously. Following this, the experimenter presented the primary stimulus object (the same Styrofoam sphere used in Hassemer & Winter, 2016) on his open hand. After removing the sphere, participants were requested to indicate both the height and the shape of the object using the index finger and thumb of a single hand (questions were asked in counter-balanced order). Crucially, we gave no instructions about what to do with respect to the dice, neither encouraging nor prohibiting moving the die out of the gesturing hand. We predict that participants are more likely to remove the die from the gesturing hand when asked the shape question, as the die would obstruct the profiled C-shape.

In Experiment 1a, we asked 185 volunteers who were pedestrians on the streets of Berlin. Of these, 114 were asked the shape question first; 71 participants were asked the height question first.[1] The interviewer then recorded whether the participant kept the die in the gesturing hand or whether and when it was moved to the other hand. In Experiment 1b, we performed a replication of this experiment with 98 beach visitors and pedestrians in the state of São Paulo (Brazil). For this conceptual replication, we used two new stimulus objects which bias against our hypothesis in that the presented physical shapes (stick and disc) exhibited exactly those form features that prior experiments had shown to influence the configuration of fingers 3-5 in the opposite direction (round curvature and extension on the vertical axis). In the stick condition, we showed participants a stick, but asked to gesture the shape of a circle with the diameter of the stick's length. In the disc condition, we showed participants a transparent disc, asking to produce a gesture

---

[1] We began the question sequence more often with the shape task because it is more informative for our research hypothesis, since it is precisely here that we expect participants to take the die out of the hand.

indicating the height of the disc. Thus, the physically presented form of the stimulus (a vertical stick, a round disc) biased against the imaginary forms[2].

2.2. Results

Out of the 185 participants in Experiment 1a, 73 people (39%) kept the dice in the respective hands for both tasks. The following analyses will not include these data because they do not speak to the research question at hand, i.e., whether the height or the shape gesture more strongly triggers the participants to remove the die. We thus focus on those participants that removed a die at least once within the experiment, either before doing the first task or before doing the second task.

For those participants who were asked the height question first, 32 out of 47 participants removed the die (68%). When they were asked the shape question following this, an additional 15 participants removed the die (32%). In contrast, when asked the shape question first, almost everybody, 57 out of 59 (97%), removed the die from the hand. The final two participants (3%) removed the die when subsequently asked the height question. A Fisher's exact test shows that the task order (height versus shape first) had a statistically reliable influence on whether the die was removed before the first or the second task ($p < 0.0001$). Thus, as predicted, the die is seen as interfering more with the shape gesture production.

There were only 6 instances in which participants put the die back into the hand after having removed it for the first task. In all of these instances, the shape question was asked first and the die was put back in for the height question. These few trials are particularly noteworthy as they suggest that participants do not avoid the dice for gesture production overall, but specifically for the shape gesture. Putting back the die for the height gesture evidences that having one die in each hand is not

---

[2] The interviewer held up the stimulus with multiple fingers and not enclosing it vertically, so as to not prime the participants by his grasp.
Online materials contain images of the stimulus objects, see https://github.com/bodowinter/new_height_shape. This repository also contains additional documentation on statistical methods (see in the results sections).

just the state that participants find themselves in when being asked to gesture (and are just unmotivated to get out of), but also the perceived normal or desirable state they come back to without being instructed. The wish to show the imaginary C-shape overrides both these reasons for keeping the die in hand.

Results of Experiment 1b abide with the results of Experiment 1a. That is, 33 participants who saw the stick and were asked to imagine a circle, put the die out of their hand (69%), compared to only 15 participants (31%) who left the die in their hand. Conversely, participants who saw the circle and were asked to imagine a line were much more likely to leave the die in the gesturing hand, namely 32 (64%) as opposed to 18 (36%) (Fisher's exact test, $p < 0.01$). The results are in accord with the imaginary form and not with the presented physical object.

2.3. Discussion

Together, these experiments serve two purposes. First, they replicate what Hassemer and Winter (2016) observed in their production study, but with tasks that showed that participants are ready to perform an unprompted action when an object interfered with presenting the imaginary C-shape in shape gestures—highlighting the role of profiling in gesture production[3]. In addition, Experiment 1b showed the same effect even though the physical stimulus biased against the targeted gesture form. Both experiments together provide converging evidence for the existence and relevance of imaginary forms in gesture production.

Given that Experiments 1a and 1b showed profiling to matter, we would expect perspective to play a role, too. This is because depending on which

---

[3] A potential concern for the production experiments 1a and 1b is that it is more difficult to hold the die in the shape task. In that case the explanation for the observed behaviors would not be based on profiling but based on physiological constraints. We believe that this is not a likely explanation because (1) it easy to hold the die between fingers 3-5 and still extend the index finger and thumb, (2) the participants who did hold the die in the hand for the shape gesture did not show any apparent difficulty of holding it in their hand, and (3) participants in Hassemer (2016) and Hassemer and Winter (2016) who gestured about a round shape but did not hold a distractor dice also sometimes had fingers 3-5 curled in, so that the physiological constraints seem to be minimal at most.

perspective a hand shape is seen from, the C-shape inside the index finger and thumb may be obstructed, or not. This idea is explored in the next experiment.

3. Experiment 2: perception experiment on perspective

3.1. Methods

For this experiment, we recruited a 3D animator (Philipp Krecklow; http://www.krecklow.net) to test the decoding process of iconicity in a perceptual task. Our stimuli are composed of a 9-stepped "pinkie curl" continuum (fingers 3-5 extended to differing degrees) where, following Hassemer and Winter (2016), we expect increases in "pinkie curl" (fingers 3-5 less extended) to result in more height as opposed to shape interpretations. We intended to replicate this finding and in addition assess the role of perspective via two different "viewing directions" and three different "viewing angles", resulting in a total of 54 (9 * 2 * 3) distinct hand stimuli as shown in Figure 2a.

For the perspectival variables, we manipulated whether the hand was seen from its ulnar side ("on pinkie" condition, rows 1-3), or its radial side ("on index" condition, rows 4-6). In the "on pinkie" condition, fingers 3-5 can actually *occlude* the profiled C-shape. In the "on index" condition, fingers 3-5 may appear within the C-shape, but they are located behind the index finger and thumb. Hence, in the "on index" condition, fingers 3-5 at most *attenuate* a clear C-shape presentation because there is less optical contrast between the index finger and fingers 3-5 than between index finger and background. As the second perspectival variable, we manipulated the viewing angle from which the hand was seen, which included three conditions: "neutral" (rows 2 and 5), 15 degrees more "from below" (rows 1 and 4), or 15° more "from above" (rows 3 and 6). This manipulates how much visual space the non-profiled fingers 3-5 take up.

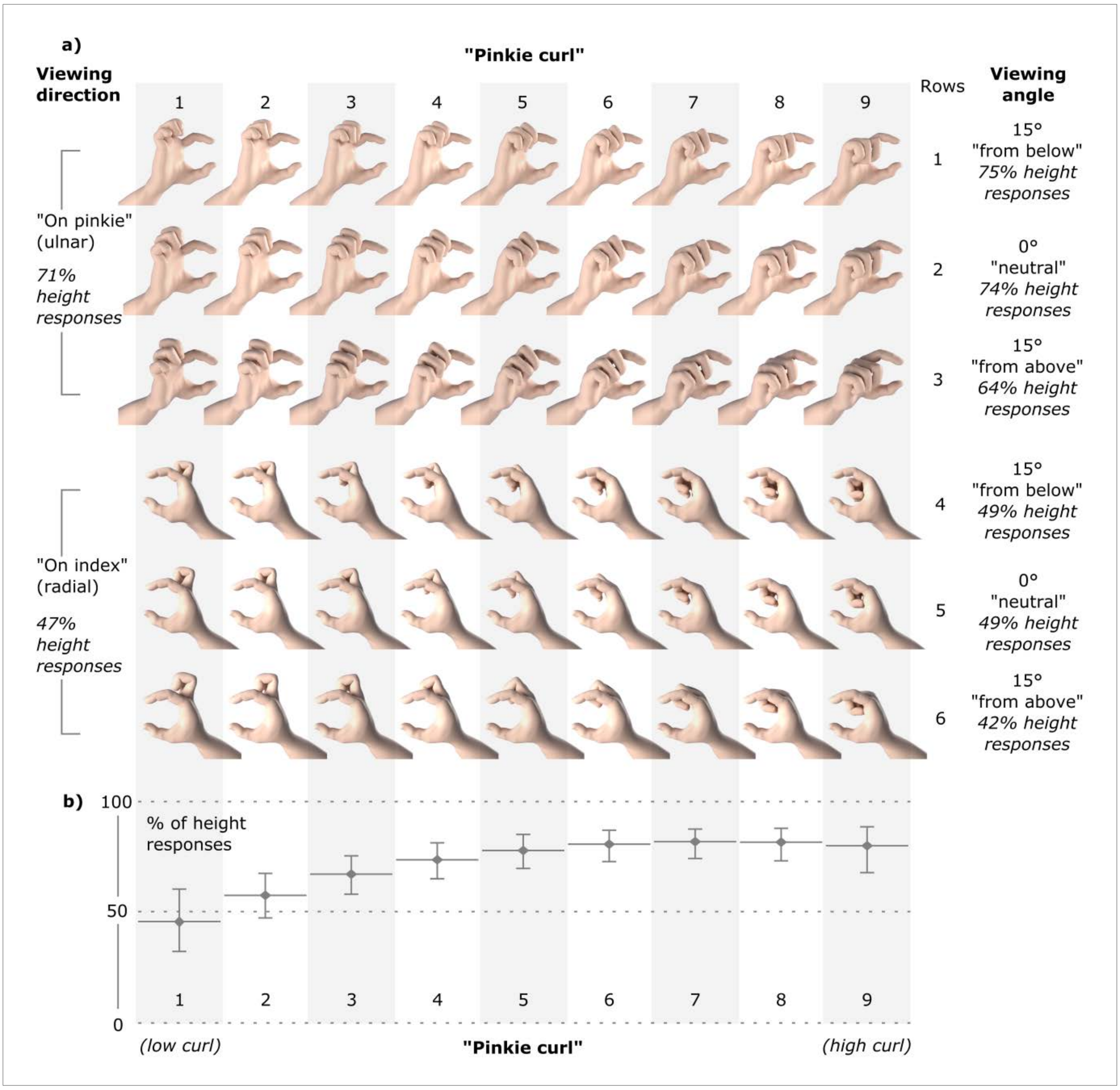

Figure 2: a) The three condition variables (9*2*3) of the 54-stimuli matrix. Columns: "Pinkie curl" continuum increasing toward the right; rows: viewing direction (rows 1-3 versus rows 4-6) and viewing angle (three rows within each row triplet) **b)** Logistic regression fits and 95% confidence intervals as a function of the "pinkie curl" continuum (averaging over viewing direction and viewing angle); 3D graphics by Philipp Krecklow (http://www.krecklow.net)

Participants were told: "On the next screen, you will be shown a picture of a hand for a few seconds. The gesture you will see characterizes an object. Please keep in mind what the hand looked like." They subsequently saw one of the handshapes from the 54-stimuli matrix for 5 seconds. After the hand disappeared, participants were asked "The gesture you just saw characterized an object. What do you think was the gesture about?", with two response options (order of options randomized): "The shape of an object" and "The height of an object". Following this, we asked an open-ended question for comments, a simple math comprehension question (4 + 17 = ?) to check whether participants paid attention to the survey, and a set of demographics questions. The experiment was managed via Qualtrics in a between-subjects design (each participant only saw one of the gestures in Figure 2). A total of 361 people were recruited via Amazon Mechanical Turk and volunteered to participate in the online experiment for 0.30 USD. After exclusion of non-native speakers and those participants that did not answer the math question correctly, there was a total of 353 participants (female = 172, male = 180, other = 1).

3.2. Results

We performed a logistic regression with the dependent measure "height/shape response" and the three predictors pinkie curl (coded as continuous factor, entered as linear and quadratic effect), viewing direction and viewing angle (*p*-values are based on likelihood ratio tests). This logistic regression model described a total of 15% of the variation in height/shape responses ($R^2 = 0.15$). There was a statistically reliable effect of pinkie curl ($\chi^2(1) = 18.66$, $p < 0.00001$). For each increase in curl by one step, the odds of observing a height response increased by 1.21 to 1 (logit coefficient: 0.19, $SE = 0.045$). The logistic regression model predicts 69% shape responses for the lowest pinkie curl (maximally extended), as opposed to only 32% for the highest pinkie curl (maximally curled in). There also was a reliable quadratic effect ($\chi^2(1) = 4.24$, $p = 0.039$), which captures the fact that at some point, further increases in pinkie curl did not lead to an increased proportion of height responses, i.e., the effect plateaus out for high curl values (see Fig. 2b). All in all, "pinkie curl" (taking the linear and quadratic effect together) described about 8% of unique variance in height/shape responses ($R^2 = 0.077$).

There was also a statistically reliable effect of viewing direction ($\chi^2(1) = 23.57$, $p < 0.0001$), which described about 8% of the variance ($R^2 = 0.08$). When fingers 3-5 were foregrounded and occluding the C-shape in the "on pinkie" condition, there were 71% height responses, as opposed to only 47% when the C-shape was in the background ("on index") (logit difference between conditions: 1.12, *SE* = 0.24). Whereas this shows a clear perspectival effect on inferred gesture form, the "viewing angle" variable showed no statistically reliable main effect ($\chi^2(1) = 2.26$, $p = 0.13$; $R^2 \approx 0$) or interaction effect with viewing direction ($\chi^2(6) = 2.26$, $p = 0.85$).

3.3. Discussion

This experiment replicates Hassemer and Winter's (2016) "pinkie curl" effect. If fingers 3-5 are curled in, a "height" interpretation is much more likely than a "shape" interpretation, consistent with the role of profiling. In addition, we found a novel effect of perspective, namely, the viewing-direction manipulation showed that one and the same hand configuration is perceived differently depending on whether the fingers 3-5 are in the fore- or background, and depending on whether they do or do not occlude the C-shape. To our knowledge, this is the first demonstration of an effect of visual perspective on the perceived *content* of representational gestures. However, we did not find an effect of viewing angle (vertical angle on the hand).

Our final experiment explores how differences in profiling lead to iconic depictions being *selective* (Clark & Gerrig, 1990; Clark, 1996), a core feature of iconicity, and a direct demonstration of the schematic and underspecified nature of gesture (Kita et al., 2017).

4. Experiment 3: perception experiment on selective depiction

4.1. Methods

The experiments so far showed how the likelihood of height/shape interpretations is affected by the position of fingers 3-5, as well as by how prominent these fingers are depending on the perspective from which the hand is viewed. A crucial element that has been implicit in this analysis so far is that height gestures specify height but underspecify shape. The complementary hypothesis is more speculative, namely that shape gestures specify shape but underspecify height (alternatively, they could equally specify height and shape). The evidence for underspecification presented here is only indirect. We showed, for example, that the obstruction of the C-shape caused participants to favor a height reading, however, this does not necessarily mean that the other aspects of the gesture are underspecified. Moreover, the choice between height and shape was enforced and only two alternatives were provided. To hone in on the notion of selective depiction, we asked participants an open question about what they take a gesture to mean, which gives them the option to specify whatever they think a gesture communicates.

Hence, in this task, we simply showed participants a "high curl" and a "low curl" stimulus from our continuum (the end points in the "on index" condition from a "neutral" angle) and subsequently asked them "What is the height of the object?" or "What is the shape of the object?", with a free text response and gave no restrictions with regard to what the response should be. 215 native American English speakers participated (recruited via Amazon Mechanical Turk; 0.30 USD reimbursement).

4.2. Results

Participants in the "high curl" condition, which profiles the index finger and thumb pad, mentioned round shapes, rectangular shapes, or both ("either round or square with flat, solid edges"), among many other responses. This highlights how this condition is compatible with multiple shapes. When asked the height question, participants often mentioned precise numerical values ("about an inch and half, no more than 2 inches") or even provided descriptions that explicitly correspond to the space between the profiled surfaces ("a few inches because that was how far apart the thumb and index fingers were"). Participants in the "low curl" condition (fingers 3-5 extended) very often gave descriptions of circular shapes (e.g., "apple", "coin", “half moon”, “perfect circle”), including descriptions that explicitly mentioned the shape profiles (“shape of a backwards letter c”). When asked the height question, participants sometimes reported heights that could not fit into the hand (“about the size of a baseball”, “25 inch diameter”), indicating height underspecification.

To quantify these observations, we coded the text responses for several features. First, whether round or rectangular objects were mentioned. The proportion of rectangular versus round shapes differed reliably by curl condition (Fisher’s exact test, p = 0.013), with 11 participants in the “high curl“ condition (11%) mentioning rectangular shapes, and 93 participants (89%) mentioning round shapes. In the “low curl“ condition, only 1 participant (1%) mentioned rectangular shapes, compared to 83 participants (99%) mentioning round shapes. Thus, despite the high overall percentage of round-shape responses, people reference comparably more rectangular objects in the “high curl” condition. That is, this gesture is seen as being less constraining when it comes to shape. On the other hand, gestures with fingers 3-5 raised are seen disproportionately more as being “about” round shapes.

We also counted vagueness markers such as “unsure”, “probably”, “hard to tell” and “of some sort”. For the shape question, 9 participants in the “high curl“ condition (8%), compared to only 1 participant in the “low curl“ condition (1%) used some form of vagueness marker, a reliable difference (Fisher’s exact test, p = 0.020). This suggests that participants were less certain about the shape implied by “high curl” gestures, consistent with the idea that these gestures underspecify shape. For the size question, we counted the proportion of responses that mentioned precise numbers, which reliably differed depending on “high” versus “low curl” (Fisher’s exact test, p = 0.0018). In the “high curl“ condition, 46 participants mentioned a precise numerical value (41%), and 66 did not (59%). On the other hand, only 21 (21%) of all participants mentioned a precise numerical value in the “low curl“ condition, as opposed to 80 (79%) who did not. Thus, it seems that the “high curl” condition attracts a larger number of responses that mention precise numerical quantities, consistent with these gestures inviting an interpretation that focuses on height.

We then counted the number of magnitude words such as “small”, “large” or “medium-sized”. Overall, the proportion of descriptions using magnitude words did not differ reliably across curl conditions (Fisher’s exact $p = 0.89$), however, there were differences with respect to *which* magnitude word was used (Fisher’s exact $p < 0.0001$). In the “high curl“ condition, not a single participant (0%) mentioned a “large” size, 5 participants (11%) mentioned a “medium” size, compared to 40 participants (89%) who mentioned a small size. In the “low curl“ condition, the responses were much more variable, with 12 “large” responses (27%), 15 “medium” responses (34%) and 17 “small” responses (39%). These results again suggest that “low curl” gestures are less specific with regard to size, as there was a bigger diversity in responses. Moreover, there was a notable absence of “large” responses in the “high curl“ condition, which is consistent with a focus on the smaller distance between the index finger and thumb pad.

4.3. Discussion

In part, these experiments provide a conceptual replication for the basic idea inherent in Experiments 1 and 2, which is that the position of the non-profiled fingers affects gesture form perception. Further, the present experiment demonstrates more clearly that iconicity is selective, which means that there is usually one piece of spatial information that, at the expense of other aspects, is in focus. In particular, a text analysis of the verbal responses shows evidence for shape-underspecification of height gestures and size-underspecification of shape gestures. For the shape question, there was an unsurprising bias to talk about round shapes regardless of which curl condition was displayed to the participant. However, the few responses that specified rectangular shapes occurred exclusively in the “high curl“ condition. Since the curvature of the index finger was kept constant, the uptake in rectangular shape responses is presumably because participants were trying to imagine which shape would fit into the empty space spanned between the index finger and thumb pad, which, following the idea of selective depiction, could be any shape. In addition, there were more verbal markers of insecurity to the shape question in the "high curl" condition. If people are less certain about what shape is implied by the gesture—which is expected to happen when a gesture underspecifies shape—, they are more likely to express that uncertainty about the shape verbally.

When asked the height question, participants were more likely to use precise numerical information in their written responses when seeing a "high curl" gesture as compared to a "low curl" gesture. This is in line with the notion that the "high curl" gesture imposes stronger constraints on a particular height, namely the distance between the index finger and thumb pad. In the "low curl" condition, which was predicted to underspecify size, we also found more variability in the mentioned sizes, including various "large" and "medium" responses. In the "high curl" condition, with the focus on the distance between the index finger and thumb pad, there were many more "small" responses. Altogether, these textual responses provide clear evidence for the selectivity of iconicity in gesture perception.

5. General discussion

Altogether, our results shed light on the cognitive processes involved in decoding iconic information from representational gestures. We replicated and extended the findings from Hassemer and Winter (2016), providing another series of empirical tests that highlight the importance of distinguishing between imagined gesture form and physical form. In addition, we showed new evidence demonstrating how the cognitive process of "profiling" mediates between physical form and gesture form. Our results clearly show that when people gesture, it is not the entirety of the gesturing hand that is equally in focus, but only specific aspects of articulators within it (see also Sowa, 2006; Hassemer, 2016). Similarly, it has been claimed for signs in signed languages, where researchers distinguish between selected and non-selected fingers (Mandel, 1981; Brentari, 1998).

Given that the decoding process from physical form to gesture form involves profiling, we predicted that the perspective from which a physical form is seen should have an influence on the perception of iconic features. We found partial evidence for this view, showing how making fingers 3-5 more prominent (and occluding the C-shape) by virtue of perspective leads to shifts in gesture form perception. The support for perspectival influences, however, is only partial, as we failed to find an effect of viewing angle (even though they spanned 30° of hand rotation). One possible explanation for the absence of an angle effect could be that participants are mentally correcting the viewing angle (compare Marr & Nishihara, 1978; see also Tarr & Pinker, 1989: 277). It is possible that mental rotation processes impact the process of decoding iconicity in ways that weaken the effect under investigation, i.e., the meaning of a gesture would not be based on the handshape in its presented rotation, but the way it was mentally rotated in object identification. This idea needs further testing.

Finally, we provided an empirical test for the idea of selective depiction (Clark & Gerrig, 1990; Clark, 1996), i.e., iconicity necessarily leaves out information and only represents particular features of a referent—those that are disclosed by how "gesture form" is perceived. As stated by Arnheim (2004 [1969]: 177), "gesture limits itself intelligently to emphasizing what matters" (see also Kita et al., 2017). This aspect of gesture form perception was targeted in Experiment 3, where free text responses showed that participants focus less on shape in "height" gestures. While the gesture form analysis of height gestures suggests that no shape information is conveyed (two small opposed surfaces can enclose objects of whatever shape), the shape gestures' C-shape within index finger and thumb can communicate a specific size, as well. Our data does not contradict the existence of size information in shape gestures, but demonstrates that overall size plays a backgrounded role for these gestures.

It may appear that the contrast we study here was categorical in that is was always either size *or* shape. However, categoricity is imposed by the binary-response experiment design and it cannot be determined from our results whether the corresponding perceptual processes involved are categorical or not (cf. Spivey, 2007). Also, the concepts of size and shape themselves are not categorical nor even mutually exclusive. A shape can be presented in a specific size thus conveying height and shape at the same time. Potentially, two gestures are perceived as belonging to two distinct categories—in accord with what Emmorey and Herzig (2003) found for ASL signs. While our experiments were not designed to test categorical perception, gesture form analyses of height and shape gestures suggests that there is no smooth transition between the shape profile of two separate opposing surfaces and one continuous C-shaped surface (see Figure 1 and Hassemer, 2016). Sevcikova Sehyr and Cormier (2016) found indications that even non-signing participants interpreted dynamically presented handling signs categorically. However, in contrast to our experiment, participants were asked to directly compare the physical characteristics of signs, whereas we asked for the *interpretation* of the gesture with respect to spatial information (gesture form). A test of categorical perception in gesture interpretation remains to be conducted.

Our results furthermore speak to the literature on the various types of gestures that are sometimes subsumed under the banner "precision grip". Precision grip gestures mark different spaces between the index finger and the thumb to refer to very different meanings, including metaphorical meanings, i.e., referring to abstract or non-spatial referents (see Kendon, 2004; Lempert, 2011; Calbris, 2011; Winter et al., 2014). These meanings are based on the underlying gesture form specifying either height or shape. Our profiling-based analysis of these gestures allows us to make predictions for which type of precision grip gesture occurs in which meaning contexts, which also has ramifications for particular metaphorical affordances. For example, we would predict that the fingers 3-5 should be curled in for gestures about "tiny numbers" (Winter et al., 2014), since these numerical gestures index metaphorical size, rather than shape characteristics.

A characteristic of all experiments in this paper is that the investigated gestures occur with little context. Gestures are inherently multifunctional (Kok et al., 2016), and gestures are furthermore embedded in discourse practices and situated in the local context (Kendon, 2004; Streeck, 2009). One and the same gesture can have different interpretations depending on which other gestures are produced in its temporal vicinity, how it is employed together with speech, and other aspects of the gesture's context. We deliberately stripped away such context to provide some insight into the raw material of gestural communication. The same way that we have learned about speech perception by presenting speech sounds in isolation (e.g., Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967), we can learn from presenting gestures in isolation to complement existing research that looks at gestures in context. We think that at the basis of any contextual modulation of gestural interpretations, there is also a gesture-only interpretation of gesture form that involves rather low-level processes, such as profiling. Moreover, from a purely methodological perspective, it is difficult and potentially problematical to investigate distinct features of gestural iconicity we discuss here (profiling, perspective, selective depiction) in naturalistic data of context-rich multimodal communication.

To finish, let us draw a parallel of our experiments to speech comprehension. There, too, language users need to map physical phenomena onto abstract phonological categories, and a lot of research in speech focuses on the interface between phonetics and phonology. In the case of gestural iconicity (in contrast to sign language, but see Liddell, 2003; Emmorey, 2003), several researchers have glossed over the distinction between physical form and gesture form, in analogy to the distinction between phonetics and phonology. The distinction was either not made, or it was made only implicitly. Our experiments offer evidence for the process of profiling, which is just one of the cognitive operations bridging the gesture's bare physical form to its gesture form (Hassemer, 2016; Hassemer & McCleary, under rev.). The evidence discussed in this paper exemplifies the need for models that break down the very broad and heterogeneous cognitive strategy of iconicity into its building blocks.

References

Arnheim, R. (2004 [1969]). *Visual Thinking*. Berkeley: University of California Press.

Brentari, D. (1998). *A Prosodic Model of Sign Language Phonology*. London: Bradford.

Calbris, G. (2011). *Elements of meaning in gesture*. Amsterdam: John Benjamins.

Cienki, A. (1998). Metaphoric gestures and some of their relations to verbal metaphoric expressions. In J.P. Koenig (Ed.), *Discourse and cognition* (pp. 189-204). Stanford, CA: CSLI Publications.

Clark, H. H. (1996). *Using Language*. Cambridge: Cambridge University Press.

Clark, H. H., & Gerrig, R. J. (1990). Quotations as demonstrations. *Language*, *66*, 764-805.

de Jorio, A. (2001). *Gesture in Naples and Gesture in Classical Antiquity: A Translation of Andrea de Jorio's La Mimica Degli Antichi Investigata Nel Gestire Napoletano* (Reprint). Indiana University Press.

Emmorey, K. (Ed.). (2003). *Perspectives on Classifier Constructions in Sign Languages*. Mahwah, NJ and London: Lawrence Erlbaum Associates.

Emmorey, K. (2014). Iconicity as structure mapping. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(20130301).

Emmorey, K., & Herzig, M. (2003). Categorical versus gradient properties of classifier constructions in ASL. In K. Emmorey (Ed.), *Perspectives on Classifier Constructions in Signed Languages* (pp. 222-246). Mahwah: Lawrence Erlbaum Associates.

Goldin-Meadow, S. (2003). *Hearing Gesture: How our Hands Help Us Think*. Cambridge, MA: Harvard University Press.

Goldin-Meadow, S., Cook, S. W., & Mitchell, Z. A. (2009). Gesturing gives children new ideas about math. *Psychological Science*, *20*, 267-272.

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science*, *12*, 516-522.

Goodwin, C. (2003). Pointing as situated practice. In S. Kita (Ed.), *Pointing: Where Language, Culture, and Cognition Meet* (pp. 217-241). Mahwah, New Jersey: Lawrence Erlbaum.

Hassemer, J. (2009). *Chief Pointing-Eye. Multiarticulatory Metaphorical Gestures Expressing High*. Unpublished MA thesis, European University Viadrina, Frankfurt (Oder).

Hassemer, J. (2016). *Towards a Theory of Gesture Form Analysis. Imaginary Forms as Part of Gesture Conceptualisation, with Empirical Support from Motion-Capture Data*. Dissertation, RWTH Aachen University.

Hassemer, J., Joue, G., Willmes, K., & Mittelberg, I. (2011). Dimensions and mechanisms of form constitution: Towards a formal description of gestures. In *Proceedings of Gesture and Speech in Interaction (GESPIN)*. Bielefeld.

Hassemer, J., & McCleary, L. (under rev.). The multidimensionality of pointing. *Gesture*. Preprint accessible under https://www.researchgate.net/publication/322570050_The_multidimensionality_of_pointing. DOI 10.13140/RG.2.2.31597.08166

Hassemer, J., & Winter, B. (2016). Producing and perceiving gestures conveying height or shape. *Gesture*, *15*, 404-424.

Jamalian, A., & Tversky, B. (2012). Gestures alter thinking About time. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 503-508). Austin, TX: Cognitive Science Society.

Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.

Kendon, A. (2014). Semiotic diversity in utterance production and the concept of 'language'. *Phil. Trans. R. Soc. B*, *369*, 20130293.

Kita, S. (2003) (Ed.). *Pointing: Where Language, Culture, and Cognition Meet*. Mahwah, New Jersey: Lawrence Erlbaum.

Kita, S., Alibali, M. W., & Chu, M. (2017). How Do Gestures Influence Thinking and Speaking? The Gesture-for-Conceptualization Hypothesis. *Psychological Review*, *124*, 245-266.

Kok, K., Bergmann, K., Cienki, A., & Kopp, S. (2016). Mapping out the multifunctionality of speakers' gestures. *Gesture*, *15*, 37-59.

Krauss, R. M., Morrel-Samuels, P., & Colasante, C. (1991). Do conversational hand gestures communicate? *Journal of Personality and Social Psychology*, *61*, 743-754.

Langacker, R. W. (2008). *Cognitive Grammar. A basic introduction*. Oxford/NewYork: Oxford University Press.

Lempert, M. (2011). Barack Obama, being sharp: Indexical order in the pragmatics of precision-grip gesture. *Gesture*, *11*, 241-270.

Lewis, T. N., & Stickles, E. (2017). Gestural modality and addressee perspective influence how we reason about time. *Cognitive Linguistics*, *28*, 45-76.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*, 431-461.

Liddell, S. K. (2003). *Grammar, Gesture, and Meaning in American Sign Language*. Cambridge: Cambridge University Press.

Mandel, M. A. (1981). *Phonotactics and Morphophonology in American Sign Language*. PhD thesis, UC Berkeley.

Mandel, M. A. (1981). Phonotactics and morphophonology in American Sign Language. PhD dissertation, University of California, Berkeley.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society B: Biological Sciences*, *200*, 269-294.

McNeill, D. (1992). *Hand and Mind*. Chicago: The University of Chicago Press.

McNeill, D. (2005). *Gesture and Thought*. Chicago: The University of Chicago Press.

Mittelberg, I., & Waugh, L. R. (2009). Metonymy first, metaphor second: A cognitive-semiotic approach to multimodal figures of thought in co-speech gesture. In C. Forceville & E. Urios-Aparisi (Eds.), *Multimodal Metaphor* (pp. 329–356). Berlin/New York: Mouton de Gruyter.

Mittelberg, I. (2014). Gestures and iconicity. In C. Müller, J. Bressem, A. Cienki, E. Fricke, S.H. Ladewig, D. McNeill & J. Bressem (Eds.), *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction* (pp. 1712-1732). Berlin: De Gruyter Mouton.

Müller, C. (1998). *Redebegleitende Gesten. Kulturgeschichte – Theorie – Sprachvergleich*. Berlin: Arno Spitz.

Müller, C. (2004). Forms and uses of the Palm Up Open Hand: A case of a gesture family? In C. Müller & R. Posner (Eds.), *The Semantics and Pragmatics of Everyday Gestures*. (pp. 233-256). Berlin: Weidler.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Sandler, W., & Lillo-Martin, D. (2006). *Sign Language and Linguistic Universals*. Cambridge: Cambridge University Press.

Sevcikova Sehyr, Z., & Cormier, K. (2016). Perceptual categorization of handling handshapes in British Sign Language. *Language & Cognition*, *8*, 501-532.

Sowa, T. (2006). *Understanding Coverbal Iconic Gestures in Shape Descriptions*. Dissertation, University of Bielefeld.

Spivey, M. (2007). *The Continuity of Mind*. Oxford: Oxford University Press.

Streeck, J. (2008). Depicting by gesture. *Gesture*, *8*, 285-301.

Streeck, J. (2009). *Gesturecraft: The Manu-facture of Meaning*. Amsterdam: John Benjamins.

Talmy, L. (2000). *Toward a cognitive semantics. Vol. I concept structuring systems*. Cambridge: MIT Press.

Talmy, L. (2018). *The targeting system of language*. Cambridge: MIT Press.

Tarr, M., & Pinker, S. (1989). Mental rotation and orientation-dependence shape Recognition. *Cognitive Psychology*, *21*, 233-282.

Wilcox, S. (2004). Cognitive iconicity: Conceptuals paces, meaning, and gesture in signed language. *Cognitive Linguistics*, *15*, 119-148.

Winter, B., Perlman, M., & Matlock, T. (2014). Using space to talk and gesture about numbers: Evidence from the TV News Archive. *Gesture*, *13*, 377-408.