counting range of one hand, suggesting that finger counting postures only activate the corresponding mental number representations when embedded in an appropriate task. Second, the absence of a *counting hand* congruency effect shows that using the non-starting hand does not necessarily activate the respective mental representation for larger numbers. Third, the finding that finger postures and numbers interact based on their respective relative sizes demonstrates a more flexible size activation through finger postures than previously assumed. This is in line with the idea of a *generalized magnitude system*, which is assumed to "encode information about the magnitudes in the external world that are used in action" (Walsh 2003, p 486). Specifically, showing almost all fingers of one hand is associated to large magnitudes and showing very few fingers to small magnitudes. The present study shows that only under certain task demands subjects activate a one-to-one correspondence between fingers and numbers. In other situations, magnitudes might not have to be exactly the same, but rather proportional to be associated.

Acknowledgments
This research is supported by DFG grant FI 1915/2-1 "Manumerical cognition".

References

Di Luca S, Pesenti M (2007) Masked priming effect with canonical finger numeral configurations. Exp Brain Res 185(1): 27–39. doi: 10.1007/s00221-007-1132-8

Domahs F, Moeller K, Huber S, Willmes K, Nuerk H–C (2010) Embodied numerosity: implicit hand-based representations influence symbolic number processing across cultures. Cognition 116(2):251–266. doi:10.1016/j.cognition.2010.05.007

Glenberg AM, Witt JK, Metcalfe J (2013) From the revolution to embodiment: 25 years of cognitive psychology. Perspect Psychol Sci 8(5):573–585. doi:10.1177/1745691613498098

Fischer MH (2008) Finger counting habits modulate spatial-numerical associations. Cortex 44(4): 386–92. doi:10.1016/j.cortex.2007.08.004

Walsh V (2003) A theory of magnitude: common cortical metrics of time, space and quantity. Trends Cogn Sci 7(11):483–488. doi: 10.1016/j.tics.2003.09.002

Object names correspond to convex entities

Rahel Sutterlütti, Simon Christoph Stein, Minija Tamosiunaite, Florentin Wörgötter
Faculty of Physics: Biophysics and Bernstein Center for Computational Neuroscience, Göttingen, Germany

Commonly one assumes that object-identification (and recognition) requires complex—innate as well as acquired—cognitive processes (Carey 2011), however, it remains unclear how objects can be individuated, segregated into parts, and identified (named) given the high degree of variability of the sensory features which arise even from similar objects (Geisler 2008). Gestalt laws, relying on shape parameters and their relations; for example edge-relations, compactness, or others; seem to play a role in this process (Spelke et al. 1993). Specifically, there exist several results from psychophysics (Hoffman and Richards 1984, Biederman 1987, Bertamini and Wagemans 2013) and machine vision (Siddiqi and Kimia 1995, Richtsfeld et al. 2012), which demonstrate that convex-concave surface transitions can be used for object partitioning.

Here we are now trying to discern to what degree such a partitioning corresponds to our language-expressible object "understanding". To this end, a total of 10 real scenes, consisting of 3D point cloud data and the corresponding RBG image, have been analyzed. Scenes were recorded by RGB-D sensors (Kinect), which provide 3D point cloud data and matched 2D RGB images. Scenes were taken from openly available machine vision data bases (Richtsfeld et al. 2012, Silberman et al. 2012). We segmented the scenes into 3D entities using convex-concave transitions in the point cloud by a model-free machine vision algorithm, the details of which are described elsewhere (LCCP Algorithm, Stein et al. 2014). This is a purely data-driven segmentation algorithm, which does not use any additional features for segmentation and works reliably for in-door RGB-D scenes with a depth range of approx. 0.5 to 5 meters using only 2 parameters to set the resolution. Note, due to the limited spatial resolution of the RGB-D sensors, small objects cannot be consistently labeled. Thus, segments smaller than 3 % of the image size were manually blackened out by us as they most often represent sensor noise. We received a total of 247 segments (i.e. about 20–30 per image). Segments are labeled on the 2D RGB image with different colors to make them distinguishable for the observer. To control for errors introduced by image acquisition and/or by the computer vision algorithm, we use the known distance error function of the Kinect sensor (Smisek et al. 2011) to calculate a reliability score for every segment.

We asked 20 subjects to compare the obtained 247 color-labeled segments with the corresponding original RGB image, asking: "*How precisely can you name it?*"; and recorded their utterances obtaining 4,940 data points. Subsequently we analyzed the utterances and divided them into three groups: 1) precise naming of a segment (e.g. "table leg"), where it does not play a role whether or not subjects would use unique names (e.g. "table leg", "leg", and "table support" are equally valid), 2) definite failure/impossibility to name a segment and 3) unclear cases, where subjects stated that they are not sure about the identification.

One example scene is shown in Fig. 1a. Using color-based segmentation (BenSalah et al. 2011) the resulting image segments rarely correspond to objects in the scene (Fig. 1b) and this is also extremely dependent on illumination. Unwanted merging or splitting of objects will, *regardless* of the chosen segmentation parameters, generically happen (e.g. "throat + face", "fridge-fragments", etc. Figure 1b). Instead of using 2D color information, here point clouds were 3D-segmented along concave/convex transitions. We observed (Fig. 1b) that subjects many times used different names (e.g. "face" or "head") to identify a segment, which are equally valid as both describe a valid conceptional entity (an object). There are however several cases where segments could not be identified. We find that on average 64 % of the segments could be identified, 30 % not, and there were 6 % unclear cases. Are these 30 % non-identified segments possibly (partially) due to machine vision errors? To assess this, we additionally considered the *reliability* of the individual segments. Due to the discretization error of the Kinect (stripy patterns in Fig. 1c), data at larger distances become quadratically more unreliable (Smisek et al. 2011) leading to merging of segments. When considering this error source, we find that subjects could more often identify reliable segments (Fig. 1e, red) and unrecognized cases dropped accordingly (green). The red lettering in Fig. 1d marks less reliable segments and, indeed, identification is lower or more ambivalent for those segments as compared to the more reliable ones.

The here performed segmentation generically renders identifiable object parts (e.g. "head", "arm", "handle" of fridge, etc.). Clearly, no purely data-driven method exists, which would allow detecting complex, *compound* objects (e.g. "woman") as this requires additional conceptual knowledge. Furthermore, we note that we are here not concerned with higher cognitive aspects, relating to context analysis, hierarchization, categorization, and other complex processes. Our main observation is that the purely geometrical (low-level) breaking up of a 3D scene, most often leads to entities for which we have an internal object or object-part concept which may

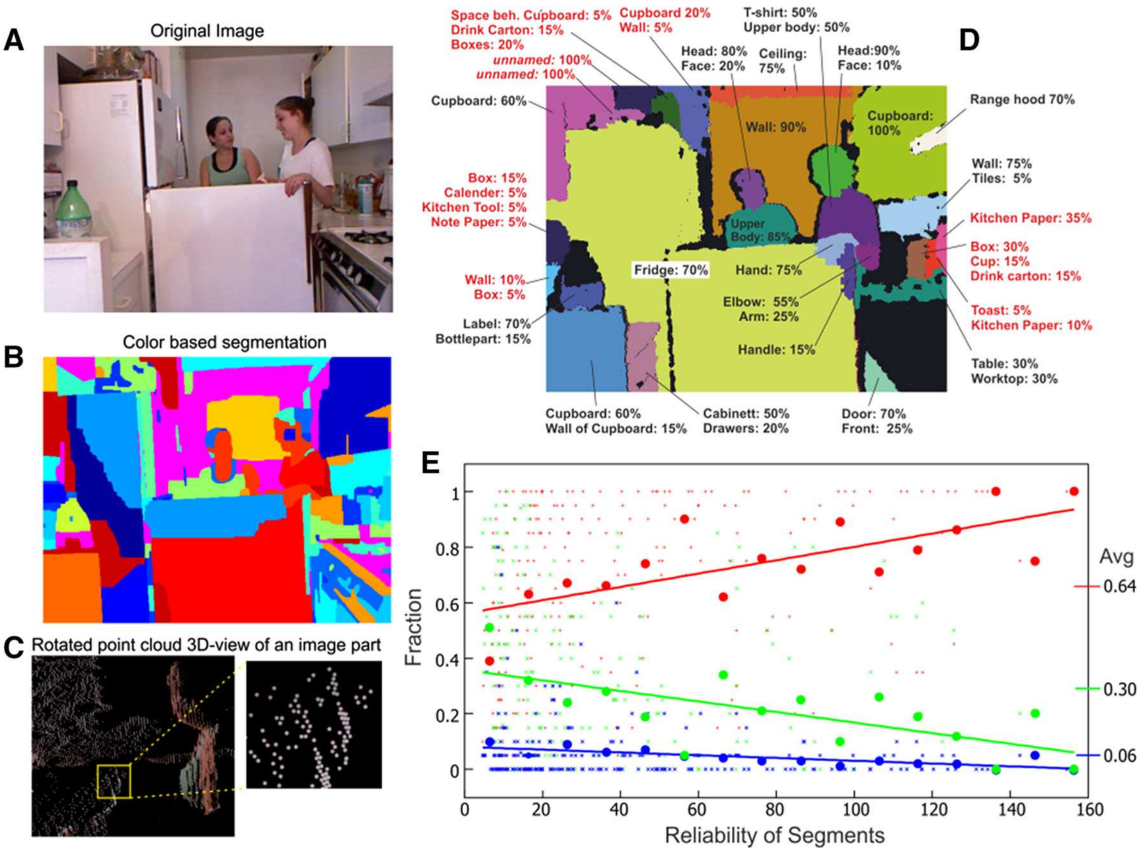

Fig. 1 Humans can with high reliability identify image segments that result from splitting images along concave-convex surface transitions. **a** One example scene used for analysis. **b** Color-based segmentation of the scene. **c** Point cloud image of parts of the scene. **d** 3D-segmented scene and segment names used by our subjects to identify objects. Missing percentages are the non-named cases. *Red* lettering indicates segments with reliability less than 50. **e** Fraction of identified (*red*), not-identified (*green*) and unclear (*blue*) segments for the complete data set plotted against their reliability. Fat dots represent averages across reliability intervals [0,10]; [10,20]; …; [150,160]. The ability to identify a segment increases with reliability. Grand averages (*red* 0.64, *green* 0.30, *blue* 0.06) for all data are shown, too

reflect the low-level perceptual grounding of the “bounded region” hypothesis formulated by Langacker (1990) as a possible foundation for grammatical entity construal.

It is known that color, texture and other such statistical image features vary widely (Geisler 2008). Thus, object individuation cannot rely on them. By contrast, here we find that convex-concave transitions between 3D-surfaces might represent the required prior to which a contiguous object concept can be unequivocally bound. These transitions render object boundaries and, consequentially leads to the situation that we can name them.

In addition, we note that this bottom-up segmentation can easily be combined with other image features (edge, color, etc.) and also—if desired—with object models where one now can go beyond object individuation towards true object recognition.

References

Ben Salah, M, Mitiche, A, Ayed, IB (2011) Multiregion image segmentation by parametric kernel graph cuts. IEEE Trans Image Proc. 20(2):545–557

Bertamini, M, Wagemans, J (2013) Processing convexity and concavity along a 2-D contour: figure-ground, structural shape, and attention. Psychon Bull Rev 20(2):191–207

Biederman I (1987) Recognition-by-components: a theory of human image understanding. Psychol Rev 94:115–147

Carey S (2011) Precis of ‘The origin of concepts’ (and commentaries), behav Brain Sci 34(3):113–167

Geisler W (2008) Visual perception and the statistical properties of natural scenes. Ann Rev Psy 59:167–192

Hoffman D, Richards W (1984) Parts of recognition. Cognition 18(13):65–96

Langacker RW (1990) Concept, image, and symbol: the cognitive basis of grammar. Mouton de Gruyter, Berlin

Richtsfeld A, Morwald T, Prankl J, Zillich M, Vincze M (2012) Segmentation of unknown objects in indoor environments. In: Proceedings of IEEE Conference on EEE/RSJ intelligent robots and systems (IROS), pp 4791–4796

Siddiqi K, Kimia BB (1995) Parts of visual form: computational aspects. IEEE Trans Pattern Anal Mach Intel 17:239–251

Silberman N, Hoiem D, Kohli P, Fergus R (2012) Indoor segmentation and support inference from RGB-D images. In: Proceedings of European conference on computer vision (ECCV), pp 746–760

Smisek J, Jancosek M, Pajdla T (2011) 3D with Kinect. In: Proceedings of international conference comp vision (ICCV), pp 1154–1160

Spelke ES, Breinlinger K, Jacobson K, Phillips A (1993) Gestalt relations and object perception: a developmental study. Perception 22(12):1483–1501

Stein S, Papon J, Schoeler M, Wörgötter F (2014) Object partitioning using local convexity. In: Proceedings of IEEE conference on computer vision and pattern recognition (CVPR), 2014. http://www.cv-foundation.org/openaccess/content_cvpr_2014/papers/Stein_Object_Partitioning_using_2014_CVPR_paper.pdf

The role of direct haptic feedback in a compensatory tracking task

Evangelia-Regkina Symeonidou, Mario Olivari, Heinrich H. Bülthoff, Lewis L. Chuang
Max Planck Institute for Biological Cybernetics, Tübingen, Germany

Haptic feedback systems can be designed to assist vehicular steering by sharing manual control with the human operator. For example, direct haptic feedback (DHF) forces, that are applied over the control device, can guide the operator towards an optimized trajectory, which he can either augment, comply with or resist according to his preferences. DHF has been shown to improve performance (Olivari et al. submitted) and increase safety (Tsoi et al. 2010). Nonetheless, the human operator may not always benefit from the haptic support system. Depending on the amount of the haptic feedback, the operator might demonstrate an over- reliance or an opposition to this haptic assistance (Forsyth and MacLean 2006). Thus, it is worthwhile to investigate how different levels of haptic assistance influence shared control performance.

The current study investigates how different gain levels of DHF influence performance in a compensatory tracking task. For this purpose, 6 participants were evenly divided into two groups according to their previous tracking experience. During the task, they had to compensate for externally induced disturbances that were visualized as the difference between a moving line and a horizontal reference standard. Briefly, participants observed how an unstable air- craft symbol, located in the middle of the screen, deviated in the roll axis from a stable artificial horizon. In order to compensate for the roll angle, participants were instructed to use the control joystick. Meanwhile, different DHF forces were presented over the control joystick for gain levels of 0, 12.5, 25, 50 and 100 %. The maximal DHF level was chosen according to the procedure described in (Olivari et al. 2014) and represents the best stable performance of skilled human operators. The participants' performance was defined as the reciprocal of the median of the root mean square error (RMSE) in each condition.

Figure 1a shows that performance improved with in- creasing DHF gain, regardless of experience levels. To evaluate the operator's contribution, relative to the DHF contribution, we calculated the ratio of overall performance to estimated DHF performance without human input. Figure 1b shows that the subject's contribution in both groups de- creased with increasing DHF up to the 50 % condition. The contribution of experienced subjects plateaued between the 50 and 100 % DHF levels. Thus, the increase in performance for the 100 % condition can mainly be attributed to the higher DHF forces alone. In contrast, the inexperienced subjects seemed to completely rely on the DHF during the 50 % condition, since the operator's contribution approximated 1. However, this changed for the 100 % DHF level. Here, the participants started to actively contribute to the task (operator's contribution >1). This change in behavior resulted in performance values similar to those of the experienced group Our findings suggest that the increase of haptic support with our DHF system does not necessarily result in over-reliance and can improve performance for both experienced and inexperienced subjects.

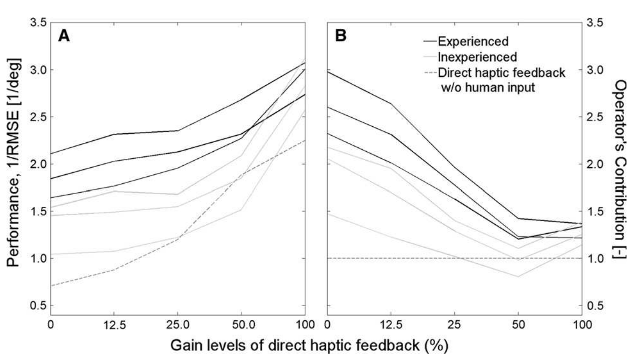

Fig. 1 **a** Performance of the experienced and in experienced participants as well as the baseline of direct haptic feedback (DHF) assistance without human input for increasing haptic gain. **b** The ratio of overall system performance to DHF performance without human input for increasing haptic gain

References

Forsyth BAC, MacLean KE (2006) Predictive haptic guidance: intelligent user assistance for the control of dynamic tasks. IEEE Trans Visual Comput Graph 12(1):103–13

Olivari M, Nieuwenhuizen FM, Bülthoff HH, Pollini L (2014) An experimental comparison of haptic and automated pilot support systems. In: AIAA modeling and simulation technologies conference, pp 1–11

Olivari M, Nieuwenhuizen F, Bülthoff H, Pollini L (submitted) Pilot adaptation to different classes of haptic aids in tracking tasks. J Guidance Control Dyn

Tsoi KK, Mulder M, Abbink DA (2010) Balancing safety and support: changing lanes with a haptic lane-keeping support system. In: 2010 IEEE international conference on systems, man and cybernetics, pp 1236–1243

Comprehending negated action(s): embodiment perspective

Nemanja Vaci[1], Jelena Radanović[2], Fernando Marmolejo-Ramos[3], Petar Milin[2,4]
[1] Alpen-Adria University Klagenfurt, Austria; [2] University of Novi Sad, Serbia; [3] University of Adelaide, Australia; [4] Eberhard Karls Universität Tübingen, Germany

Keywords
Embodied cognition, Negation, Mental simulation, Sentence comprehension

According to the embodied cognition framework, comprehension of language involves activation of the same sensorimotor areas of the brain that are activated when entities and events described by language structures (e.g., words, sentences) are actually experienced (Barsalou 1999). Previous work on the comprehension of sentences showed support for this proposal. For example, Glenberg and Kaschak (2002) observed that judgment about sensibility of a sentence was facilitated when there was congruence between the direction of an action implied by the sentence and the direction of a movement required for making a response, while incongruence led to slower responses. It was also shown that linguistic markers (e.g., negation) could modulate mental simulation of concepts (Kaup 2001). This finding was explained by the two-step negation processing: (1) a reader simulates a sentence as if there is no negation; (2) she negates the simulated content to reach full meaning. However, when a negated action was announced in preceding text, negated clause was processed as fast as the affirmative one (Lüdtke and Kaup 2006). The mentioned results suggest the mechanism of negation processing can be altered contextually.

In this study, we aimed at further investigating the effects of linguistic markers, following the assumptions of embodied