ijbd

International Journal of Behavioral Development
2014, Vol. 38(6) 481–486

DOI: 10.1177/0165025414544233
ijbd.sagepub.com
SAGE

# Universal ontology: Attentive tracking of objects and substances across languages and over development

Trix Cacchione,[1] Marcello Indino,[1] Kazuo Fujita,[2] Shoji Itakura,[2] Toyomi Matsuno,[2] Simone Schaub,[1] and Federica Amici[3]

**Abstract**

Previous research has demonstrated that adults are successful at visually tracking rigidly moving items, but experience great difficulties when tracking substance-like "pouring" items. Using a comparative approach, we investigated whether the presence/absence of the grammatical count–mass distinction influences adults and children's ability to attentively track objects versus substances. More specifically, we aimed to explore whether the higher success at tracking rigid over substance-like items appears universally or whether speakers of classifier languages (like Japanese, not marking the object–substance distinction) are advantaged at tracking substances as compared to speakers of non-classifier languages (like Swiss German, marking the object–substance distinction). Our results supported the idea that language has no effect on low-level cognitive processes such as the attentive visual processing of objects and substances. We concluded arguing that the tendency to prioritize objects is universal and independent of specific characteristics of the language spoken.

**Keywords**

cross cultural differences, multiple-object tracking, linguistic relativity, object–substance distinction

The ontological distinction between objects and substances is one of the most basic in conceptual development (Prasada, Ferenz, & Haskell, 2002). Solid objects are conceived of as individuated entities which are defined by shape (e.g., a chair, a car, a duck). In contrast, substances are conceived of as non-solid, non-individuated entities which are not defined by shape (e.g., sand, powder, gel; see Prasada, 1999).

Children and infants younger than 1 year can differentiate between objects and substances (Hall, 1996; Hespos, Ferry, & Rips, 2009; Huntley-Fenner, Carey, & Solimando, 2002; Imai & Gentner, 1997; Samuelson & Smith, 1999; Soja, 1992; Soja, Carey, & Spelke, 1991, 1992), suggesting that the fundamental ontological distinction between objects and substances is available to children even before language is acquired (Landau, Jones, & Smith, 1992; Soja, 1992; Soja et al., 1991, 1992). Infants conceive the surrounding world as being populated by objects, whose unity and boundaries they perceive through spatio-temporal information operating on specific core principles (e.g., Spelke, 1994, 2000; Spelke, Breinlinger, Macomber, & Jacobson, 1992; Spelke & Kinzler, 2007). The core principle of cohesion (i.e., objects move as solid bound features) is proposed to be the most powerful principle defining the ontological category of objects (Bloom, 2000; Pinker, 1997; Scholl, 2007). The ability to spot out "object-like" features is thus of elementary relevance in conceptual development, and may naturally entail a habitual focus on objects as the relevant features in the surroundings. Indeed, individual solid objects are the prioritized units of perceptual and cognitive processing such as physical reasoning, categorizing or enumerating, not only in infants (e.g., Carey & Xu, 2001; Leslie, Xu, Tremoulet, & Scholl, 1998; Scholl, 2007; Spelke, 1994), but throughout human lifespan (for reviews, see Bloom, 2000; Carey & Xu, 2001; Scholl, 2001), and possibly in our phylogenetic development (Cacchione & Call, 2010; Cacchione, Hrubesch, & Call, 2013; Mahajan, Barnes, Blanco, & Santos, 2009).

Crucial evidence of the fundamental role of the object/substance core distinction in cognitive processing has been found in studies investigating the tracking and quantifying of solid objects and non-solid substances. Tracking and quantifying substances (e.g., water or sand) is difficult because substances constantly change their shape and split into parts while moving. Non-solid substances cannot therefore be individuated, as they cannot be separated neither through spatio-temporal (e.g., internal coherence) nor featural criteria (e.g., stable shape), unless a special criterion is provided (e.g., a *pile* of sand, a *cup* of milk). Recent research has demonstrated adults' and infants' difficulties to keep track of individuated portions of non-solid substances. In particular, when confronted with several targets moving on a screen among similar distractors, adults could not track them as well when a substance-like pouring motion was performed (vanMarle & Scholl, 2003). Similarly, 8–12-month-old infants could not track or process substances as well as solid objects (Cheries, Mitroff, Wynn, & Scholl, 2008; Chiang & Wynn, 2000; Huntley-Fenner et al., 2002; but see also Cacchione, 2013; Hespos, Dora, Rips, & Christie, 2012; vanMarle & Wynn,

[1] University of Zurich, Switzerland
[2] Kyoto University, Japan
[3] Max Planck Institute for Evolutionary Anthropology, Germany

**Corresponding author:**
Federica Amici, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz, 6, Leipzig 04107, Germany.
Email: abbepu@yahoo.it

2011). Taken together, these studies suggest that the ontological distinction between objects and substances is a universal core property in human cognitive processing.

Some authors, however, have proposed that this object/substance distinction is not part of the core cognition, but rather acquired through language (e.g., Quine, 1960, 1969). In non-classifier languages, like English or German, count nouns allowing the plural marker (e.g., bottle, bowl, castle) are used to label individuated bounded objects, while mass nouns (e.g., water, sugar, sand) are used to label substances, which can be only indirectly individuated by putting them into portions (e.g., *a bottle* of water or *piles* of sand; Imai & Gentner, 1997; Prasada et al., 2002; Soja et al., 1991; Subrahmanyam, Landau, & Gelman, 1999). According to Quine (1960), the ontological distinction between objects and substances would not be present in children's mental life before they master the linguistics of the count–mass distinction. Speakers of classifier languages, like Japanese, which lack the count–mass distinction, would therefore not acquire the object/substance distinction in the way speakers of non-classifier languages do.

The effect of language on cognitive processes is however controversial (see Evans & Levinson, 2009; Wolff & Holmes, 2011). Some studies, for example, suggest that even basic perceptual cognitive processes can be affected by the language spoken (e.g., Berent, Steriade, Lennertz, & Vaknin, 2007, on perception of acoustic signals). Moreover, language can partially affect cognitive processes like categorization (e.g., Davidoff, Davies, & Roberson, 1999; Regier & Kay, 2009, on color categorization; Bowerman & Choi, 2001, on spatial categories). Under certain circumstances, speakers of non-classifier languages categorize objects attending to shape rather than material, possibly because objects are individuated on the basis of shape; in contrast, speakers of classifier languages are biased to focus on material, probably because their nouns do not explicitly individuate objects (Imai & Gentner, 1997; Imai & Mazuka, 2003, 2007; Li, Dunham, & Carey, 2009; J. Lucy & Gaskins, 2001; J. A. Lucy, 1992). Although the ability to think about fixed regular-shaped objects as individuated entities is universal (e.g., Li et al., 2009), the characteristics of individual languages might affect which particular perceptual dimensions the speakers pay attention to, leading them to spontaneously focus more on shape or material and thus partially shaping their concepts of objects and substances (Imai & Gentner, 1997; J. A. Lucy, 1992; Saalbach & Imai, 2012).

Is it therefore possible that speaking a classifier/non-classifier language might affect participants' ability to track objects and substances? Multiple object tracking involves the simultaneous tracking of multiple moving targets among distractors and measures cognitive processing by tapping on perception, attention and working memory (Bahrami, 2003; Pylyshyn & Storm, 1988; Sears & Pylyshyn, 2000; Trick, Jasper-Fayer, & Sethi, 2005) Previous research suggests that the ability to track objects and substances relies on mid-level processes (i.e., implying both perceptual and conceptual processing, Scholl, 2001), and that it is sensitive to the object–substance distinction (vanMarle & Scholl, 2003). When confronted with several moving objects/substances on a screen, for example, English-speakers are considerably more successful at tracking rigidly moving objects as compared to unbound substances (Scholl & Pylyshyn, 1999; Scholl, Pylyshyn, & Feldman, 2001; vanMarle & Scholl, 2003).

In this study, we used a language-free multiple object tracking task modeled after vanMarle & Scholl (2003) to investigate whether the ability to better track rigidly moving objects over unbound substances (vanMarle & Scholl, 2003) is really a cognitive universal, or whether the grammatical count–mass distinction affects adults and children's ability to visually process objects/substances. In their study, VanMarle and Scholl (2003) tested all speakers of a non-classifier language, which by imposing a habitual focus on shape and cohesiveness through the count–mass distinction might explain participants' relative failure at tracking unbound substances. If the count–mass distinction really affected participants' ability to track objects/substances, speakers of non-classifier languages should be more successful at tracking objects as compared to substances, while speakers of a classifier language should perform more similarly when tracking objects and substances. On the other hand, if language has no influence on basic cognitive processes, like the ability to process objects/substances, speakers of both types of languages should show identical processing behavior.

## Methods

### Participants

Participants were 48 adults and 48 children from two different linguistic populations, speaking either a classifier language (Japanese) or a non-classifier language (Swiss German). The 24 native Japanese-speaking adults (12 females, 12 males; mean age = 23 years, age range = 19–30 years) and the 24 native Japanese-speaking children (10 females, 14 males; mean age = 5 years 9 months, age range = 5 years–6 years 7 months) were tested in testing facilities at the University of Kyoto. The 24 native Swiss German-speaking adults (14 females, 10 males; mean age = 27 years, age range = 20–49 years) were tested in testing facilities at the University of Zurich, and the 24 native Swiss German-speaking children (10 females, 14 males; mean age = 6 years 6 months, age range = 4 years 7 months–6 years 8 months) in a quiet room in their kindergartens in the greater Zurich area. All children were monolinguals. A few native Japanese-speaking adults could also speak some very basic English, while Swiss German-speaking adults had no knowledge of a classifier language (having learned French and secondarily English at school; two languages that do not differ from Swiss German in terms of count–mass distinction). All participants had normal or corrected-to-normal acuity and were rewarded for their participation with a small gift. All participants were randomly assigned to one of the two experimental conditions (object condition, substance condition), resulting in an equal distribution of participants in terms of language (i.e., Japanese, Swiss German) and age class (i.e., adults, children) per condition.

### Apparatus and stimuli

The tracking animations were constructed with Adobe Flash CS4 and were presented on a portable computer (Acer TravelMate 4672LMi, 1024 MB RAM, Intel Core Duo T2300, 1.66 GHz). The built-in TFT LCD-monitor accessed an ATI Mobility Radeon X1400 graphic card and had a refresh rate of 60 hertz. The 15-inch monitor was 30.5 cm wide and 22.8 cm high and had a screen resolution of 1400 × 1050 pixels (i.e., each virtually quadratic pixel had a side-length of approximately 0.21 mm).

A tracking animation for adults consisted of ten items (five targets and five distractors) moving for a duration of 13 s, thereby changing their moving direction three times, and sequentially reaching three different intermediate landmarks. Items were 10

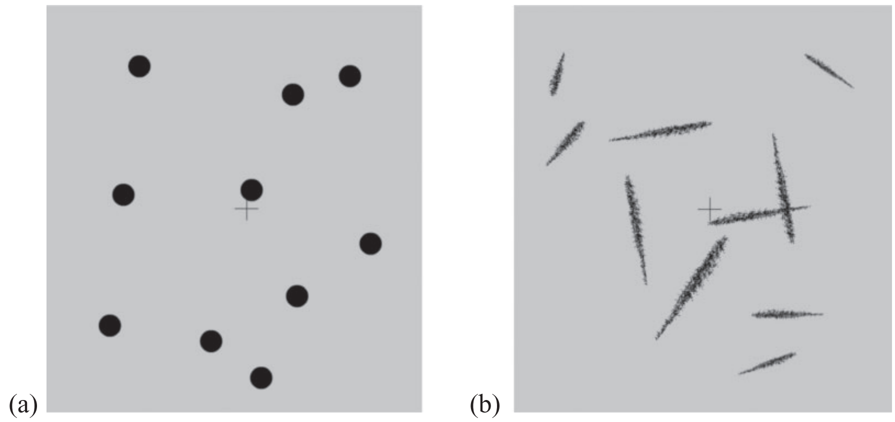


**Figure 1.** In the object condition, the items moved as rigid circles from one landmark to another (a). In the substance condition, the items moved from one landmark to the other after a substance-like pouring motion was performed (b). In both conditions, adults followed 5 targets out of 10 items, while children followed 3 targets out of 6 items.

black circles (RGB 0, 0, 0), with a diameter of 1.6 cm, subtending approximately 1.5° of visual angle. Items were presented on a grey (RGB 205, 205, 205) background (22.8 cm × 22.8 cm) positioned in the middle of the monitor (the residual display was achromatic). Initial and final position of each item and of the three intermediate landmarks were randomly defined with the constraints that a) initial and final position of items had to be in different quadrants of the background, and b) items could not overlap. The movement of the single items was therefore independent from each other, creating a non-continuous and unpredictable global moving pattern.

Tracking animations for children were scaled down versions showing only six items (three targets, three distractors) moving for 8 s. Moreover, animations for children included only one intermediate landmark. In all other respects, apparatus and stimuli were identical for adults and children.

Participants sat at a distance of about 60 cm from the monitor, so that the display subtended a visual angle of approximately 30° breadthways by 20° in height. A continuously adjustable holder permitted setting the centre of the monitor at the participants' eye level, independently of their height. The participants were instructed to continuously focus on the fixation cross at the middle of the screen. Because head movement was not necessary to capture the moving stimuli, no head restraint was employed.

### Design and procedure

As in the original task of vanMarle and Scholl (2003) animations including two kinds of movement were presented. The rigid movement in the object condition corresponded to classical multiple-object tracking displays where the items maintain their shape and boundaries while moving across the screen (Figure 1a; see Video 1 in the Supplementary material). In contrast, animations in the substance condition presented participants with "substance-like" motion (Figure 1b; see Video 2 in the Supplementary material).

In the substance condition, items only appeared as bound circles at the single static positions (initial, final, and intermediate landmark positions). Between these positions, the items "poured" as a substance-like stream of single elements. To achieve this substance-like pouring impression, the circles progressively disintegrated into a big number of individually moving pixels, progressively scaling down their diameters and pouring in clusters from one landmark to another. After reaching the new position/landmark, the pixel stream progressively merged into a circle again, progressively scaling up the diameter of the respective circle until reaching the original size. Directly after a circle reached its original diameter at its new position, the decomposition process restarted and the growing stream of pixels moved to another landmark. In their maximal dispersion (i.e., in the middle of their movement), the substance-like streams were different in magnitude, and could reach an approximate length of 6.5 cm (6°) to 8 cm (7.5°) and an approximate width of 0.5 cm (0.5°) to 1.5 cm (1.5°). The motion in the object and substance conditions followed the identical trajectories at the same speeds.

Participants were randomly assigned to one of the two conditions (between-design). After a brief instruction, two training trials were presented to familiarize participants with the task, followed by 20 different experimental trials per condition presented in a random order. The single trials were manually initiated by the experimenter. As soon as each trial started, the centre of the scene was marked by a black fixation cross having the same span as the items. Participants were requested to focus at the fixation cross and not to follow the single items with their gaze. At the beginning of each trial, half of the items indicated their status as target by blinking 8 times in 2 seconds. Afterward, all the items started to move linearly, but at slightly different speeds, from their initial positions to the first landmark. There, they changed their direction and continued their motion to a sequence of 4 further landmarks until finally reaching their static end-position. During each trial, participants were requested to intentionally track the five items defined as targets and to indicate them at the end of the animation by pointing to them at the end of each trial. Their response was not subject to any time limit. The experimenter noted the participants' responses without providing any feedback. This procedure was identical for adults and children with the exception that children had to track and indicate only three targets out of six items (see Videos 3 and 4 in the Supplementary material).

## Results

### Adults

Tracking accuracy was recorded per each trial. The mean percentage of correct tracking per condition and language population is

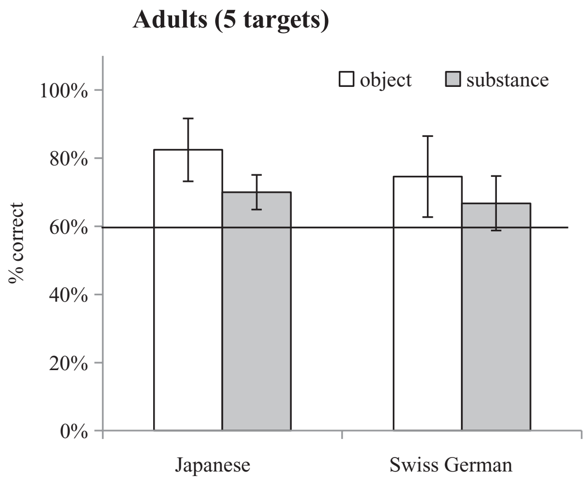


**Figure 2.** For each linguistic population (Japanese, Swiss German), mean percentage of correct responses (± SE) by adults tracking 5 target stimuli in the object and substance conditions. The horizontal line represents chance performance (= 60%, see text).

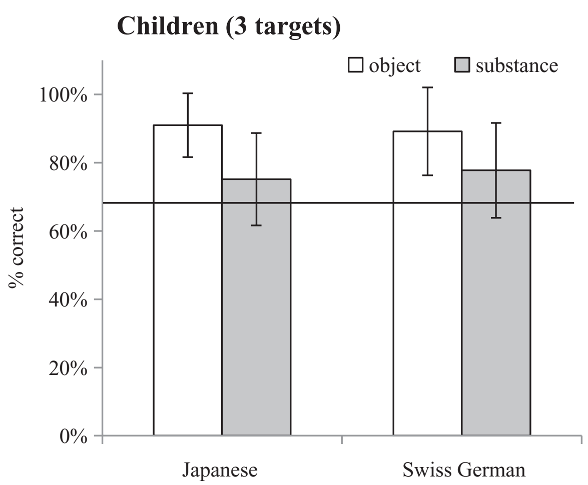


**Figure 3.** For each linguistic population (Japanese, Swiss German), mean percentage of correct responses (± SE) by children tracking 3 target stimuli in the object and substance conditions. The horizontal line represents chance performance (= 66.67%, see text).

summarized in Figure 2. In both language populations, tracking accuracy was relatively high in the object condition (Japanese, JP: 82.42%, $SD = 9.20$; Swiss German, SG: 74.58%, $SD = 11.87$) and lower in the substance condition (JP: 70.00%, $SD = 5.10$; SG: 66.75%, $SD = 8.01$). Chance performance was calculated as the performance obtained by tracking only one item and guessing for the other items ($\frac{1}{2} \times [(1/n) + 1]$); see Scholl et al., 2001; vanMarle & Scholl, 2003). It was therefore 60% for the adult subsamples who tracked 5 targets. Both populations performed above chance level (with $\alpha$ level being set at 0.050) in both conditions (in the object condition, JP: $t(11) = 8.44$, $p < .001$; SG: $t(11) = 4.26$, $p = .001$; in the substance condition, JP: $t(11) = 6.79$, $p < .001$; SG: $t(11) = 2.92$, $p = .010$).

Preliminary analyses showed no effect of order and sex. An analysis of variance (ANOVA) on tracking accuracy with condition (object vs. substance) and language population (Japanese vs. Swiss German) as between subject variables revealed a main effect of condition, $F(1, 44) = 15.660$, $p < .001$, $\eta_p^2 = .262$, with overall higher tracking rates in the object (78.50%, $SD = 11.13$) than in the substance condition (68.37%, $SD = 6.77$). Furthermore, there was a main effect of language population, $F(1, 44) = 4.691$, $p = .036$, $\eta_p^2 = .096$, with Japanese-speakers (76.20%, $SD = 9.65$) showing generally higher tracking accuracy than Swiss German-speakers (70.67%, $SD = 10.68$). Finally, the interaction of condition $\times$ language population was not significant, $F(1, 44) = 0.799$, $p = .376$, $\eta_p^2 = .018$.

### Children

The mean percentage of correct tracking per condition and language population is summarized in Figure 3. In both language populations, tracking accuracy was high in the object condition (JP: 90.97%, $SD = 9.33$; SG: 89.17%, $SD = 12.86$) and lower in the substance condition (JP: 75.14%, $SD = 13.51$; SG: 77.78%, $SD = 13.90$). Chance performance was calculated to be 66.67% (see above). Children always performed above chance level in the object condition, JP: $t(11) = 9.02$, $p < .001$; SG: $t(11) = 6.06$, $p < .001$, but only Swiss-German-speakers in the substance condition, JP: $t(11) = 2.17$, $p = .053$; SG: $t(11) = 2.77$, $p = .018$.

Preliminary analyses showed no effect of order and sex. An ANOVA on tracking accuracy with condition and language population as between subject variables revealed a main effect of condition, $F(1, 44) = 14.157$, $p < .001$, $\eta_p^2 = .243$, with overall higher tracking rates in the object (90.07%, $SD = 11.03$) than in the substance condition (76.25%, $SD = 13.47$). There was no effect of language population, $F(1, 44) = .013$, $p = .909$, $\eta_p^2 = .000$, and the interaction of condition $\times$ language population was not significant, $F(1, 44) = .377$, $p = .542$, $\eta_p^2 = .009$.

## Discussion

Regardless of the language spoken, both adults and children were better at tracking objects than substances. These results are in line with previous findings (e.g., vanMarle & Scholl, 2003) and support the core principles view according to which humans are universally endowed with basic principles defining the ontological category of objects, laying the base for humans' attentional focus on objects in perceptual and cognitive processing (Bloom, 2000; Pinker, 1997; Scholl, 2007; Spelke, 1994). In line with other studies evidencing a prioritized focus on objects across lifespan (Carey & Xu, 2001; Scholl, 2001) and in non-human primates (Cacchione & Call, 2010; Cacchione et al., 2013; Mahajan et al., 2009), these results strengthen the view that the object/substance distinction is conceptually mediated, universal across languages and probably partly innate.

In our study, adult speakers of a classifier language (Japanese) and of a non-classifier language (Swiss German) successfully tracked rigidly moving objects and substance-like pouring items, but were more successful at tracking objects. Despite important linguistic differences, therefore, speakers of a classifier language conceptually treat objects/substances like speakers of a non-classifier language. Children performance was similar to that of adults. Regardless of the language spoken, all children were better at tracking objects than substances, and this was especially true for Japanese children, who performed at chance level only when tracking substances. The fact that only Japanese children performed at chance level when tracking substances further confirms that speaking a classifier language provides no advantage at visually processing substances. To our knowledge, the present study was the first one assessing children's tracking of objects versus substances with a standard multiple-object tracking task (Scholl & Pylyshyn, 1999; Scholl et al., 2001; Sears & Pylyshyn, 2000; vanMarle & Scholl, 2003), adding on to the comparatively small body of research using standard multiple-object tracking in children (O'Hearn, Hoffman, & Landau, 2010; O'Hearn, Landau & Hoffman, 2005; Trick et al., 2005).

Our results clearly confirm the view that universal ontologies drive cognitive processing across all languages. Moreover, investigating visual processing of objects/substances, we did not find the same blend of universal commonalities and cross-linguistic differences as others did (Gao & Malt, 2009; Imai & Gentner, 1997; Saalbach & Imai, 2007, 2012; Zhang & Schmitt, 1998). A possible explanation for these differences is the level at which language might affect cognitive processes. If universal ontologies and language specific effects on cognition really coexist, language effects are probably second-order, shaping the precedent universal structures. The universal object/substance distinction might therefore be viewed as primary, manifesting itself early in ontogenetic development and cognitive processing, as hardwired low-level constraint. Language effects might be instead secondary, emerging later in development and being limited to higher level cognitive processing (e.g., categorization, inductive reasoning). Such perspectives would be in line with the view that the relation between language and cognitive processing is bidirectional, and linguistic structures both reflect and modify universal cognitive structures (Imai & Mazuka, 2007; Roberson & Hanley, 2010; Saalbach & Imai, 2012).

Finally, Japanese adults (but not children) were generally better than Swiss Germans at tracking both objects and substances. This might be due to several factors (e.g., higher commitment or higher willingness to perform correctly: e.g., Muramoto, Yamaguchi, & Kim, 2009), but does not change the key finding that tracking moving objects appears universally easier than tracking moving substances.

So far, we might therefore conclude that language does not affect low-level cognitive processes such as the attentive visual processing of objects and substances. Instead, adults and children across different language populations adhere equally to the object/substance distinction and focus on objects as the prioritized unit of cognitive processing. The tendency to prioritize objects thus appears universal, irrespectively of whether the language spoken syntactically differentiates objects from non-discrete substances. This supports the view that cognitive processing is mediated by universal and probably innate constraints, adding to a substantial body of findings rejecting the idea of a strong determining and unidirectional influence of language on cognitive processing.

## Funding

This research was supported by a grant from the Swiss National Science Foundation (SNSF) to the first author (ID Number PZ00P1_126320/1) and was written while the last author held a Humboldt Research Fellowship for Postdoctoral Researchers (Humboldt ID Number 1138999).

## Supplemental material

Supplemental videos are available online at http://jbd.sagepub.com/supplemental.

## References

Bahrami, B. (2003). Object property encoding and change blindness in multiple object tracking. *Visual Cognition*, *10*, 949–963.

Berent, I., Steriade, D., Lennertz, T., & Vaknin, V. (2007). What we know about what we have never heard: Evidence from perceptual illusions. *Cognition*, *104*, 591–630.

Bloom, P. (2000). *How children learn the meanings of words*. Cambridge, MA: MIT Press.

Bowerman, M., & Choi, S. (2001). Shaping meanings for language: Universal and language-specific in the acquisition of spatial semantic categories. In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 475–511). Cambridge, UK: Cambridge University Press.

Cacchione, T. (2013). The foundations of object permanence: Does perceived cohesion determine infants' appreciation of the continuous existence of material objects. *Cognition*, *128*, 397–406.

Cacchione, T., & Call, J. (2010). Do gorillas (*Gorilla gorilla*) and orangutans (*Pongo pygmaeus*) fail to represent objects in the context of cohesion violations? *Cognition*, *116*, 193–203.

Cacchione, T., Hrubesch, C., & Call, J. (2013). Chimpanzees (*Pan troglodytes*) and bonobos (*Pan paniscus*) quantify split solid objects. *Animal Cognition*, *16*, 1–10.

Carey, S., & Xu, F. (2001). Infants' knowledge of objects: Beyond object files and object tracking. *Cognition*, *80*, 179–213.

Cheries, E. W., Mitroff, S. R., Wynn, K., & Scholl, B. J. (2008). Cohesion as a constraint on object persistence in infancy. *Developmental Science*, *11*, 427–432.

Chiang, W. C., & Wynn, K. (2000). Infants' representation and tracking of multiple objects. *Cognition*, *77*, 169–195.

Davidoff, J., Davies, I., & Roberson, D. (1999). Colour categories in a stone-age tribe. *Nature*, *398*, 2003–204.

Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, *32*, 429–492.

Gao, Y., & Malt, B. C. (2009). Mental representation and cognitive consequences of Chinese individual classifiers. *Language and Cognitive Processes*, *24*, 1124–1179.

Hall, G. (1996). Naming solids and nonsolids: Children's default construals. *Cognitive Development*, *11*, 229–264.

Hespos, S. J., Dora, B., Rips, L. J., & Christie, S. (2012). Infants make quantity discriminations for substances. *Child Development*, *83*, 554–567.

Hespos, S. J., Ferry, A., & Rips, L. (2009). Five-month-old infants have different expectations for solids and liquids. *Psychological Science*, *20*, 603–611.

Huntley-Fenner, G., Carey, S., & Solimando, A. (2002). Objects are individuals but stuff doesn't count: Perceived rigidity and cohesiveness influence infants' representations of small groups of distinct entities. *Cognition*, *85*, 203–221.

Imai, M., & Gentner, D. (1997). A cross-linguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition*, *62*, 169–200.

Imai, M., & Mazuka, R. (2003) Re-evaluation of linguistic relativity: Language-specific categories and the role of universal ontological knowledge in the construal of individuation. In: D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the issues*

*of language and thought* (pp. 430–646). Cambridge, MA: MIT Press.
Imai, M., & Mazuka, R. (2007). Language-relative construal of individuation constrained by universal ontology: Revisiting language universals and linguistic relativity. *Cognitive Science*, *31*, 385–413.
Landau, B., Jones, S., & Smith, L. (1992). Perception, ontology, and naming in young children: Commentary on Soja, Carey, and Spelke. *Cognition*, *43*, 85–91.
Leslie, A. M., Xu, F., Tremoulet, P. D., & Scholl, B. J. (1998). Indexing and the object concept: Developing "what" and "where" systems. *Trends in Cognitive Sciences*, *2*, 10–18.
Li, P., Dunham, Y., & Carey, S. (2009). Of substance: The nature of language effects on entity construal. *Cognitive Psychology*, *58*, 487–524.
Lucy, J., & Gaskins, S. (2001). Grammatical categories and the development of classification preferences: A comparative approach. In: M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 257–283). Cambridge, UK: Cambridge University Press.
Lucy, J. A. (1992). *Grammatical categories and cognition: A case study of the linguistic relativity hypothesis*. Cambridge, UK: Cambridge University Press.
Mahajan, N., Barnes, J. L., Blanco, M., & Santos, L. R. (2009). Enumeration of objects and substances in non-human primates: Experiments with brown lemurs (*Eulemur fulvus*). *Developmental Science*, *12*, 920–928.
Muramoto, Y., Yamaguchi, S., & Kim, U. (2009). Perception of achievement attribution in individual and group contexts: Comparative analysis of Japanese, Korean, and Asian-American results. *Asian Journal of Social Psychology*, *12*, 199–210.
O'Hearn, K., Hoffman, J. E., & Landau, B. (2010). Developmental profiles for multiple object tracking and spatial memory: Typically developing preschoolers and people with Williams syndrome. *Developmental Science*, *13*, 430–440.
O'Hearn, K., Landau, B., & Hoffman, J. E. (2005). Multiple object tracking in people with Williams syndrome and in normally developing children. *Psychological Science*, *16*, 905–912.
Pinker, S. (1997). Words and rules in the human brain. *Nature*, *387*, 547–548.
Prasada, S. (1999). Names for things and stuff: An Aristotelian perspective. In: R. Jackendoff, P. Bloom & K. Wynn (Eds.), *Language, logic, and concepts* (pp. 119–146). Cambridge, MA: MIT Press.
Prasada, S., Ferenz, K., & Haskell, T. (2002). Conceiving of entities as objects and as stuff. *Cognition*, *83*, 141–165.
Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, *3*, 179–197.
Quine, W.V. (1960). *Word and object*. Cambridge, MA: MIT Press.
Quine, W.V. (1969). *Ontological relativity and other essays*. New York, NY: Columbia University Press.
Regier, T., & Kay, P. (2009). Language, thought, and color: Whorf was half right. *Trends in Cognitive Sciences*, *13*, 439–446.
Roberson, D., & Hanley, R. (2010). Relatively speaking. An account of the relationship between language and thought in the color domain. In B. Malt & P. Wolff (Eds.), *Words and the mind: How words capture human experience* (pp. 183–198). New York, NY: Oxford University Press.
Saalbach, H., & Imai, M. (2007). The scope of linguistic influence: Does a classifier system alter object concepts? *Journal of Experimental Psychology: General*, *136*, 485–501.
Saalbach, H., & Imai, M. (2012). The relation between linguistic categories and cognition: The case of numeral classifiers. *Language and Cognitive Processes*, *27*, 381–428.
Samuelson, L. K., & Smith, L. B. (1999). Early noun vocabularies: Do ontology, category structure and syntax correspond? *Cognition*, *73*, 1–33.
Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition*, *80*, 1–46.
Scholl, B. J. (2007). Object persistence in philosophy and psychology. *Mind & Language*, *22*, 563–591.
Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology*, *38*, 259–290.
Scholl, B. J., Pylyshyn, Z. W., & Feldman, J. (2001). What is a visual object? Evidence from target merging in multiple-object tracking. *Cognition*, *80*, 159–177.
Sears, C. R., & Pylyshyn, Z. W. (2000). Multiple object tracking and attentional processing. *Canadian Journal of Experimental Psychology*, *54*, 1–14.
Soja, N. N. (1992). Inferences about the meanings of nouns: The relationship between perception and syntax. *Cognitive Development*, *7*, 29–45.
Soja, N. N., Carey, S., & Spelke, E. S. (1991). Ontological categories guide young children's inductions of word meaning: Object terms and substance terms. *Cognition*, *38*, 179–211.
Soja, N. N., Carey, S., & Spelke, E. (1992). Discussion: Perception, ontology, and word meaning. *Cognition*, *45*, 101–107.
Spelke, E. S. (1994). Initial knowledge: Six suggestions. *Cognition*, *50*, 431–445.
Spelke, E. S. (2000). Core knowledge. *American Psychologist*, *55*, 1233–1243.
Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, *99*, 605–632.
Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science*, *10*, 89–96.
Subrahmanyam, K., Landau, B., & Gelman, R. (1999). Shape, material, and syntax: Interacting forces in children's learning in novel words for objects and substances. *Language and Cognitive Processes*, *14*, 249–281.
Trick, L. M., Jaspers-Fayer, F., & Sethi, N. (2005). Multiple-object tracking in children: The "Catch the Spies" task. *Cognitive Development*, *20*, 373–387.
vanMarle, K., & Scholl, B. J. (2003). Attentive tracking of objects versus substances. *Psychological Science*, *14*, 498–504.
vanMarle, K., & Wynn, K. (2011). Tracking and quantifying objects and non-cohesive substances. *Developmental Science*, *14*, 502–515.
Wolff, P., & Holmes, K. J. (2011). Linguistic relativity. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*, 253–265.
Zhang, S., & Schmitt, B. (1998). Language-dependent classification: The mental representation of classifiers in cognition, memory, and ad evaluations. *Journal of Experimental Psychology: Applied*, *4*, 375–385.