Title:

How similar are objects and events?

Authors:

Alexis Wellwood **[corresponding]**
Department of Philosophy
wellwood@usc.edu

Susan J. Hespos
Department of Psychology
hespos@northwestern.edu

Lance Rips
Department of Psychology
rips@northwestern.edu

Keywords:

mass/count, telicity, ontology, number, psycholinguistics

Abstract:

Semanticists often assume an ontology for natural language that includes not only ordinary objects, but also events, as well as further distinctions. We link this ontology to how speakers represent static and dynamic entities. Specifically, we test how speakers determine whether an entity counts as "atomic" by using count vs. mass (e.g., *some gleebs*, *some gleeb*) and distributive vs. non-distributive descriptions (e.g., *gleeb every second or so*, *gleeb around a little*). We then seek evidence for atomic representation in a non-linguistic task. Ultimately we suggest that natural language ontology reveals properties of language-independent conceptualization.

How similar are objects and events?

1. Introduction

Many semanticists now assume an ontology for natural language that includes not only ordinary objects like cups and tables, but also events like jumpings and fallings (e.g., Vendler 1957; Davidson 1967). A noun like *cup* applies to the object type, while a verb like *jump* applies to the event type. At the same time, developmental and cognitive psychologists have become increasingly more interested in understanding how dynamic entities are represented, and philosophers have debated whether events form a distinct metaphysical category, and, if so, with what properties (Zacks & Tversky 2001 and Casati & Varzi 2008, respectively, provide recent overviews). This paper contributes an attempt to link the semanticists' notion of an event to the psychologists' notion. We do this by investigating the conditions under which people carve up their experience into static or dynamic categories with different formal properties.

In fact, the linguistic ontology is often thought to be populated not only by objects and events, but by other categories of spatial and temporal entity including at least substances and processes (Vendler 1957; see Champollion & Krifka 2016). The sheer variety of types of entities assumed can quickly lead to conflicts, though, between what language seems to commit its speakers to, and what an austere metaphysics might prefer (Pietroski 2015). Consider the case of a solid gold ring, which can be felicitously referred to in a context either as *the gold* or *the ring*. Intuition suggests that the ring and the gold are the same thing; meanwhile, we can truthfully ascribe properties to the ring (e.g., being ornate) that are not true of the gold (Link 1983). The same can be observed for happenings: suppose that we intuit that whatever Ann did, it satisfies both *her crossing the Channel* and *her swimming*—the trouble is, we may want to say that while the crossing was fast, her swimming was slow (cf. Davidson 1985; Krifka 1989).

Making these ontological distinctions matters for a semantic theory that aims to specify the connections between what words and sentences mean, and the things that they truthfully apply to. Yet, as we've hinted, positing such distinctions in the ontology is also puzzling, if the relevant things have to be "out there" in the world. The standard interpretation of the enterprise allows the study of natural language to be a useful probe for reasoning about metaphysics; if so, semanticists will have to actively commit themselves to a world with very different properties than might be suggested by other branches of science. As the number of types of entities in the model multiply, and if two or more of those entities have to occupy the same region of space-time, it becomes more and more difficult to see how a mind-independent world putatively consisting of atoms, quarks, and strings or whatever can support the requisite metaphysics.

On the other hand, generative linguists going back at least to Chomsky (1965) have demanded that a theory of meaning relate linguistic expressions not to the world, but to non-linguistic cognition. In light of its presumed metaphysical commitments, the project of the formal semanticist is thus distinct from the goals of generative linguistics. This difference in the stated explanatory domain has led some linguists to claim that a semantic theory is not properly a part of linguistic theory at all (see Chomsky 1995; cf. Partee 1979). Meanwhile, formal semantics has continued to accrue many descriptive successes, and the question of its foundations has for the most part languished. In recent years, though, there has been an upsurge in new attempts to 'naturalize' semantics, and thus align its agenda more closely with that of generative syntax (e.g., Pietroski et al. 2009; Pietroski 2010; Lidz et al. 2011; Vogel et al. 2014). These views characterize formal semantic description as a source of explicit hypotheses about how syntactic representations align with representations in non-linguistic cognition.

This theoretical tension has led some thinkers to suppose that it is not in the purview of semantic theory to negotiate between these two opposing positions. Bach (1986a) takes an agnostic stance for what he calls "natural language metaphysics" (see also Pelletier 2011; Bach & Chao 2012). On this view, the semanticist needn't, and in fact shouldn't, commit to saying more than: the entities we posit are those that speakers talk "as if" there are (cf. Montague 1973, n. 8). Despite the presumed connection between meaning and truth in natural language semantics, and the goal of that theory to illuminate the nature of linguistic meaning, the explanatory goals of semantics on Bach's conception nevertheless exclude saying whether the entities it posits are really "out there", in some sense, or whether they are actually indicative of deeper cognitive facts about speakers.

This paper aims to get beyond "talk *as if*". We examine how, within a certain domain, semantic theory can be used to generate specific hypotheses about how speakers represent and reason about the world. The case we consider is Bach's (1986b) analogy between objects and substances, on the one hand, and events and processes on the other. We build on the results of Wellwood et al. (2016), who found evidence for these categories in conceptual categorization as a function of the 'naturalness' of novel spatial and temporal entities. There, we measured speakers' preference for describing these novel displays using count versus mass syntax (e.g., *some gleeb(ing)* vs. *some gleebs*), and found that the locations—rather than, say, the number—of boundaries in both the static images and the animations were highly predictive of speakers' preferences. In light of the semantic requirements of count syntax, we interpreted this result as evidence for formal parallels between object and event representation outside of language.

That work leaves two important questions open, and which our experiments are designed to investigate: (i) whether the observed parallel preferences between the spatial and temporal domains were due to object and event individuation per se, as opposed to something specifically limited to speakers' understanding of the semantics of nominal plurality; and (ii) whether this type of individuation exists independent of the pressures of a forced-choice linguistic task. In §2, we first discuss the linguistic evidence for the relevant ontological distinctions, their formal characterization, and previous research informing our questions. In §3, we test the hypothesis that distributive adverbials involve the same sort of event individuation as pluralized deverbal nouns (e.g., *do some gleebs*), as expected by Rothstein's (1995) analysis. In §4, we test the hypothesis that these preference tasks reveal how people conceptualize static and dynamic entities independent of language.

2. Conceptualizing the ontology

We use language to talk about all sorts of things, and many of the differences between those things don't play any particularly important role in linguistic description. For instance, no grammatical rules are sensitive to the fact that cows are a distinct kind of animal from horses, and the count nouns *cow* and *horse* share the same grammatical distribution—they can be pluralized (*cows*, *horses*), modified by numerals (*three cows*, *seventeen horses*), distributively quantified (*each cow*, *every horse*), etc. But some of the properties of the things we talk about plausibly do relate to important linguistic distinctions: mass nouns that name substances (e.g., *flesh*, *steel*) distribute differently from nouns that name objects. That is, it is difficult to predict or locate the meaning of many pluralized substance-denoting nouns (?*fleshes*, ?*steels*), to say how they should be counted (?*three flesh/es*, ?*seventeen steel/s*), or to quantify them distributively (?*each flesh*, ?*each steel*), etc.[1]

In what follows, we first review the linguistic evidence for an ontology that distinguishes, as the above examples suggest, objects from substances, as well as events from processes. The primary theoretical distinction that we target is the notion that expressions that apply to objects and events refer *atomically*, while those that apply to substances and processes refer *non-atomically*. Next, we show how the semantic analysis of the relevant morphosyntactic cues (e.g., the plural morpheme, distributive adverbials, etc.) bundles in the atomicity requirement. Finally, we combine the semanticists' notion of 'Atom' with recent results from cognitive psychology. This allows us to generate hypotheses about how participants' inferences about certain features of static and dynamic entities should interact with those linguistic features, and we put these hypotheses to the test in §3 and §4.

2.1 A précis of the linguistic evidence

Formal semanticists have distinguished objects and substances from events and processes, based on how different expressions interact with linguistic individuation, counting, and distributive quantification. Beginning with expressions that are thought to apply to objects (*cocktail*) and events (*party*), the obvious difference is that objects *exist* while events *happen*—a difference borne out in asymmetries in the naturalness of predications invoking weights or start times, (1) (cf. Casati & Varzi 2008).

(1) a. The cocktail weighed a pound.
b. ? The cocktail started at midnight.

(2) a. ? The party weighed a ton.
b. The party started at midnight.

Apart from this, there are a number of similarities. Objects and events are the kinds of things that can be paired one to one, as the intuitive understandings of the constructions in (3) illustrate (cf. Boolos 1981; Rothstein 1995). For (3a) to be judged true, it must be the case that each individual cocktail is paired with an individual shot in the relevant context. Similarly, (3b) requires that each drinking event is paired with a getting ornery event.

[1] See Gillon (1992, 2012) and Chierchia (1998) for extensive discussion of the grammatical mass/count distinction in English and other languages.

(3) a. For every cocktail you drink, you get a free shot.
b. Every time you drink, you get ornery.

Expressions that apply to objects and events can comfortably be pluralized, either using the nominal plural morpheme, or by adverbials indicating frequentative action (see Henderson 2013 and references therein for discussion over devices available in other languages). The sentences in (4) imply that Ann likes or drinks the things of the specified sort but not, for example, some arbitrary quantity of the stuff of which they're composed. In the same way, the sentences in (5) imply that Ann likes or attended events of the specified sort, not merely some arbitrary quantity of the activities that go on during them.

(4) a. Ann likes cocktails.
b. Ann drinks shots.

(5) a. Ann likes parties.
b. Ann went partying again and again.

Because object- and event-denoting expressions label individuated things, these expressions can be modified by numerals in the nominal domain, (6), as well as the verbal domain, (7).

(6) Ann had three cocktails.

(7) a. Ann went to three parties.
b. Ann went partying three times last week.

The contrasting categories—substance and process—are disanalogous from each other in exactly the same way as objects and events, and analogous in ways exactly opposite to these categories. Substances (*goo*) are physical entities, and so it makes sense to say how much they weigh, (8a), but it is not easily said when they 'start', (8b). Processes (*fun*) are temporal entities, and so can't be weighed, (9a), but it is easy say when they begin, (9b).

(8) a. The goo weighed a pound.
b. ? The goo started at midnight.

(9) a. ? The fun weighed a ton.
b. The fun started at midnight.

Substances and processes do not present themselves as obviously individuated. This claim helps to explain why expressions that apply to these categories resist composition with constructions implying one-to-one pairing, (10), or a plurality of individuated entities, (11)-(12). If such expressions do not necessarily refer to individuated entities, this also helps account for why it is not intuitively obvious what it would mean to say they occur in a certain number, (13)-(14).

(10) a. ? For every goo you eat, you get water.
b. ? Every time you see, you breathe.

(11) a. ? Ann likes goos.

b. ? Ann drank gins.

(12) a. ? Ann likes funs.

b. ? Ann saw again and again.

(13) Ann bought three goo(s).

(14) a. ? Ann had three fun(s).

b. ? Ann saw three times.

So far, we have just revisited common observations about the intuitive differences that certain nouns and verbs have with respect to individuation, and the concomitant effects on how we can use these expressions to talk about or to quantify entities. Mereological semantic approaches capture these effects as interactions between the domains of application for the respective lexical items, and the semantic commitments imposed by the morphosyntactic context.

2.2 Formalizing the analogy

A pluralized noun phrase, in almost all cases, carries the commitment that its satisfiers are non-overlapping, countable things.[2] In simple cases, like with *toys* and *spaceships*, those satisfiers are things which do not have arbitrary proper subparts in the privileged sense (i.e., in the sense of "unstructured parthood" assumed in formal semantics; see Champollion 2017). Another way of putting it is that they do not have parts that are the same type as the whole. One way of guaranteeing this commitment is to say that plural morphology on an NP presupposes that its NP denotes a set of atoms. Atomicity can be defined for entities simpliciter if (15) holds, or it can be defined relative to predicates or concepts as in (16) (cf. Krifka 1989; Grimm 2012; Rips & Hespos 2015 call this property "stratification").

(15) $\text{Atom}(x) = \neg\exists y\ (y < x)$
'*x* is an atom just in case *x* has no proper parts'

(16) $\text{Atomic}(P) = \ \forall x(P(x) \rightarrow \neg\exists y(y < x \wedge P(y)))$
'*P* is atomic just in case every *x* which is *P* has no *y*-parts that are also *P*.'

An expression that applies atomically thus provides just the sort of condition required for counting to be consistent in the first place (see Koslicki's 1997 discussion of Frege's criteria): the sorts of things they apply to are not such that you can just keep dividing them and still find numerically further instances of the same thing.[3, 4]

---

[2] Exceptions include so-called mass plurals like *suds* and *mashed potatoes*, for speakers who accept these expressions in contexts like, "How much mashed potatoes do you want?". See Ojeda (2005) and Acquaviva (2008) for discussion. For exceptions to the non-overlap requirement, see Rothstein (2010).

[3] The notion of 'dividing' is not without difficulties; see especially Zucchi & White (2001). Rothstein (2010) among others advocate a contextual parameter that fixes what counts as an atom in a context.

[4] We understand 'countability' in this sense to apply to the extralinguistic entities which a given expression may or may not apply to. This notion of countability is thus independent from that employed when talking about, for example, the distributional properties of predicates, wherein a predicate may be called 'countable' if it combines directly with cardinal number words, etc. See Rothstein (2017) for discussion of countability in the latter sense.

In Link's (1983) model (and that of many others since), plural morphology merely introduces the algebraic closure of a set, notated *. Supposing that the denotation of a count noun like *toy* is $\{a, b, c\}$, where each of *a, b,* and *c* meet the condition in (16), then **toy* would denote the set $\{a, b, c, a\oplus b, a\oplus c, b\oplus c, a\oplus b\oplus c\}$, including the atoms as well as all of the possible sums consisting of just those atoms. The denotation of a mass noun like *flesh* is also a set, but which needn't have minimal parts (i.e., it is *anti-atomic*; see Gillon 2012 for recent discussion of different approaches to the semantics of mass nouns).[5] If the plural morpheme requires that its N denote atomic objects, this can help explain why it is that languages exist in which such nouns must first be 'singularized' before they can be pluralized (Mathieu 2012; cf. Borer 2005). Wellwood (2014) interprets this as effecting a mapping from a non-atomic (or anti-atomic) domain to an atomic one, allowing a substance-denoting noun to meet the atomicity condition of the plural morpheme.

Of interest to us is not whether known lexical items satisfy an atomicity condition, but whether people's comprehension of their world is such that it traffics in such notions like 'atomicity' as it is defined in semantics. If so, we would expect participants to understand a novel noun (or verb) in just these terms, and in the same way, if the novel item is embedded in any linguistic context that requires atomicity. Such a finding would suggest the potential for semantic theory to inform how people represent and reason about the world.

What sorts of other linguistic contexts have been suggested to involve the atomicity property? Many authors have suggested that the verbal parallel of the mass/count distinction in the nominal domain is the atelic/telic distinction (e.g., Mourelatos 1978; Hoepelman & Rohrer 1980), which describes how grammar interacts with different categories of eventuality. The intuition is that telic predicates—which include information about when an eventuality is 'completed'—express properties of entities that are bounded in time in much the same way that count nouns express properties of entities bounded in space, while atelic predicates express properties of activities or processes. On Bach's (1986b) formalization, paralleling that offered by Link (1983) for objects and substances, telic VPs express properties of atomic events, whereas atelic VPs apply to non-atomic processes (though compare Rothstein 2004).

Of particular interest to us is the fact that we can identify linguistic contexts that insist that their VP be atomic, understood in a precise way as the eventive equivalent of (15)-(16). Rothstein (1995) presents just such a case: that of distributive adverbials like in *every time he walked into the room, she turned to stare*. Distributive adverbials like this express that a one-to-one mapping holds between the event described in the adverbial, and that expressed by the embedded sentence. In light of analyses like these, the sentence in (17a) should impose parallel commitments on the novel verb *gleeb* as the sentence in (17b) imposes on the novel noun.

(17) a. The star gleebed every second or so.

b. There are some gleebs.

---

[5] An alternative explanation to that implied in the text for the fact that characteristically mass nouns don't easily pluralize relates to the hypothesis that substance domains are natively closed under sum formation. Applying a closure-under-sum operation would thus be redundant (Champollion & Krifka 2016). This explanation appears to miss the fact that, when a substance-denoting phrase is pluralized, its interpretation is akin to that of a plural object-denoting noun; *fleshes* implies that there is a plurality of countable things, each of which is constituted by some flesh. It doesn't mean what *flesh* means.

2.3 Previous experimental results

Objects and events that are easy to individuate invite the use of count syntax and counting, even when these items are unfamiliar. For example, Prasada et al. (2002) found that people chose a novel noun in count syntax (*There's a blicket in the tray*) over mass syntax (*There's blicket in the tray*) to describe unfamiliar cut-out shapes when the shapes appeared non-arbitrary. In that experiment, the degree of nonarbitrariness varied as a function of the shape's repetition, regularity, or shape-relevant use.

Likewise, events that take place within a short time-span—punctual events, such as clapping or jumping—encourage comparison by number of occurrences over comparison along other dimensions, such as total duration. Barner et al. (2008) asked participants to decided who "did more Ns?" or who "did more N-ing?" for nouns derived from either durative verbs (e.g., *dance*, *run*) or punctual verbs (e.g., *jump, kick*). The participants' choice was between a character who did more repetitions (e.g., a larger number of dances or jumps) or a character whose action had a larger quantity of some other relevant dimension (usually duration, distance, or intensity). Barner et al. found that, for punctual verbs, number of repetitions dominated the alternative dimension in both syntactic contexts. For example, participants chose the character who did more jumps in answer to both "Who did more jumps?" and "Who did more jumping?" But for durative verbs, repetition mattered for count syntax but the alternative dimension for mass syntax. For "Who did more dances?" number of repetitions mattered most, but for "Who did more dancing?" duration mattered most.

[Insert figure 1 around here]

We can interpret these results as suggesting that perceived atomicity of an entity influences whether we can use count syntax to describe it, and whether counting is an appropriate dimension of comparison. For example, the regularity of a shape implies that the shape has a unity and discreteness that allows it to be counted, and similarly for the punctuality of an event. This commonality between the domains of objects and events is in line with the analogy that we sketched in §2.1 and §2.2.

To obtain more direct evidence for this correspondence, we asked participants in an earlier study (Wellwood et al. 2016) to decide between count or mass descriptions of matched images and animations. In the image condition, participants saw images with parts of equivalent size and shape at natural points (i.e., several "petals" of the same size) or images with parts divided at arbitrary points. Figure 1 provides examples of these types for images containing four and five parts. After viewing an image, participants decided whether to label it with a novel noun in count syntax (*There were some gorps*) or in mass syntax (*There was some gorp*). In the animation condition, participants saw animations of a star moving along a path, either pausing at a salient, equally-spaced point (i.e., at the cusp of the paths around several invisible petals) or at arbitrary points. Figure 2 shows a schematic view of these animation types (the flower-shaped outline was invisible in the actual displays). Participants again chose between count and mass descriptors (*The star did some gleebs* vs. *The star did some gleebing*). In both conditions, participants preferred count syntax for the naturally divided items and mass syntax for the arbitrarily divided ones. Figure 3 plots the results of the four main conditions in terms of the proportion of times participants chose count over mass phrases, and it shows a parallel difference in the two domains.

[Insert Figure 2 around here]

[Insert Figure 3 around here]

These results indicate that intuitively natural shapes and paths tend to satisfy atomicity, thus promoting the use of count syntax to describe images and animations. We suspect, however, that atomicity has a more central role to play in organizing our concepts than just encouraging the use of count syntax. First, atomicity may determine the appropriateness of verbal expressions, not just nominal ones. As we've already suggested, atomicity in the domain of eventualities supports telic descriptions (e.g., whether the event is explicitly said to start or stop or to occur a specified number of times) over atelic ones. We therefore predict that a dynamic activity with natural breaks (e.g., those in the top row in Figure 2) would encourage people to encode the scene as involving *V-ing every second or two* in preference to *V-ing around a little*, for a novel verb *V*. An activity with arbitrary breaks, however, should produce the opposite preference. Second, because atomicity makes counting possible, we would expect that people's judgments of the similarity between two groups of objects or events would be more sensitive to the cardinality of the groups for non-arbitrary than for arbitrary items. The experiments we report below test these predictions.

In what follows, we describe the static and dynamic scenes as having natural as opposed to unnatural divisions. The natural divisions create whole "petals" (images) or "loops" (animations), and so, intuitively, encode more than merely non-arbitrary division. Non-arbitrary division would seem to require a weaker cut, such that each part had the same shape, however arbitrary that shape might appear. The sense of "natural" we intend, then, is in the sense of "carved at the joints"—the location of the divisions is important in addition to the overall similarity of shape (cf. Prasada et al. 2002). In ongoing research, we explicitly test whether naturalness in this sense or mere non-arbitrariness is sufficient to garner the kinds of results reviewed in this section and reported below.

3. Experiment 1: Verbal language preference

This experiment makes use of participants' preferences in pairing dynamic displays with English sentences containing the novel verb *gleeb*, as a probe of whether they categorize what's happening in a scene as an event or a process. If our participants' knowledge of English can be captured roughly as we described in §2.1, then they should know that how *gleeb* should be understood differs depending on the morphosyntactic frame it appears in. Wellwood et al. (2016) observed that participants preferred to pair dynamic displays whose movement pattern was broken up at regular intervals with sentences containing a pluralized novel noun *gleebs* (cf. Prasada et al. 2002; Barner et al. 2008), in line with this general expectation.

However, it is possible that the semantic commitments of the nominal plural morpheme drove the effect observed by Wellwood et al. (2016), rather than event individuation per se. It is uncontroversial that adding *-s* to a noun leads to an interpretation in terms of pluralities, each of which is constituted by some bounded entity. Confronted with dynamic displays, it is possible that participants merely supposed that such a condition was better fulfilled by the display with pauses at regular intervals than at irregular intervals, irrespective of deploying the category 'event'. Even if we wanted to say that this supposition nonetheless reveals access to that category, it is possible that the nominal context is 'special' in inviting reference to such entities, and that these data aren't particularly informative for understanding the semantics of verbs.

In light of these questions, participants in the present experiment viewed a dynamic display, and judged whether the scene was better described using a sentence with a novel verb and either (i) a distributive adverbial or (ii) a non-distributive adverbial. The semantics of distributivity require Atoms, in the technical sense described above, which is theoretically one and the same property ultimately required by the plural morpheme. On one trial, for example, a participant might see an object (in our case, a star) traverse an invisible petal-shaped path with five petals, which pauses in its movement only at the center of the 'flower'. Following this, they are asked to choose whether they would prefer a distributively-quantified or a non–distributively-quantified sentence to describe the scene. On another trial, the participant might see the star traverse the same overall path, but pause in its movement at five arbitrary points along the path, and be asked to make the same preference judgment. In line with previous results linking categorization and language, we expected that participants would be able to detect the differences between these two types of displays, and if these differences are significant, it should lead them to make different linguistic selections at test.

Participants. We recruited 48 participants through the Northwestern University Department of Linguistics subject pool. They received 1 lab credit for 1 hour of participation. The present study took around 15 minutes of each participants' time, and the remaining time was used for other studies.

Method. The experiment method is represented schematically in Figure 4. Participants first see an animation with *n* natural or unnatural divisions presented on the screen. Immediately following this, a screen with the question, "How would you prefer to describe that animation?" appeared, along with the two options listed in (18). Participants indicated their preference by pressing 'f' or 'j' on the keyboard.

(18) a. The star gleebed every second or so.
b. The star gleebed around a little.

Design. We manipulated the factor DIVISION (natural, unnatural) and NUMBER (of divisions: 4, 5, 6, 7, 8, 9). Our stimuli were presented blocked by DIVISION, in counter-balanced order. On each trial, participants were asked to judge how they would prefer to describe the animation, given the choices in (18).

Materials. We used the same animation stimuli as Experiment 1 in Wellwood et al. (2016), shown schematically in Figure 2. Each combination of the factors DIVISION and NUMBER delivered a total of 12 unique animations, each of which took 3 seconds on average to complete.

[Insert Figure 4 around here]

Results and discussion. Our participants strongly preferred to label the naturally-divided animations with the novel verb *gleeb* when it was paired with the distributive modifier *every second or so*, and the opposite preference was observed when the animations were unnaturally-divided, as shown in Figure 5. Furthermore, there was no overall effect of the number of divisions in a display. As we discuss below, however, we did find a small interaction effect: the preference for distributive syntax increased somewhat along with the number of divisions for the naturally-divided animations, but remained fairly flat for the unnaturally-divided animations.

[Insert Figure 5 around here]

For our statistical analysis, we conducted logistic mixed effects regressions (LMERs) with DIVISION and NUMBER as predictors, including both random intercepts and slopes (i.e., the maximal model m, Barr et al. 2013). The dependent variable was the proportion of distributive choices. The χ^2 and p values that we report for a given effect were derived by conducting an ANOVA between a model m that includes a given predictor φ, and the model m' that is exactly like m except φ is not included as a predictor. If the difference between these two models is significant, this suggests that φ plays a significant role in explaining the data.

Participants strongly preferred to pair the naturally-divided animations with the distributively-modified sentence (*The star gleebed every second or so*), and the unnaturally-divided animations with the non-distributive sentence (*The star gleebed around a little*) (proportion distributive choices: natural .81, unnatural .17). This effect was borne out in the statistical analysis as a strong main effect of DIVISIONS, $\chi^2(1) = 50.1$, $p < .0001$. This effect was as we predicted: the observation by Wellwood et al. (2016) that participants strongly preferred to pair count syntax with naturally-divided animations extends to a preference for distributive modification for those displays, suggesting that the difference in conceptualization of the two, rather than the syntax, drives these effects.

Participants' preference for distributive modification did not generally change as the number of divisions increased, as the difference in proportion of those choices was relatively similar even for the maximally different numbers of divisions (e.g., 4-divisions: .48, 9-divisions: .52). This was revealed in the lack of a main effect of NUMBER, $\chi^2(1) = 1$, $p = .31$ This result would be unexpected if the mere number of divisions highlighted the temporal structure of the event, but is unsurprising if conceptualization rather than low-level perceptual features drives the linguistic preference.

That being said, we did observe that participants' preference for distributive syntax in the naturally-divided animations increased along with the number of divisions, unlike for the unnaturally-divided animations (increase in proportion of distributive syntax between animations with 4 and 9 divisions: natural 7 percentage points, unnatural 2 percentage points). This effect was revealed in an interaction between the factors DIVISION and NUMBER, $\chi^2(1) = 4.6$, $p = .032$. We did not necessarily predict this result; however, it is compatible with our main claims in the following way: naturally-divided animations suggest categorization in terms of events, which (we contend) are entities that privilege counting. Conceptualizing the display in this way, the number dimension could be made more salient, and thus lead participants to the heightened perception that such displays are more suitable for distributive quantification.

This last result hints that we may be able to detect effects of atomicity on people's conception of entities, even in a task that does not directly rely on language labeling. If naturally-divided displays, like those in Figures 1 and 2, evoke atomic objects and events, then we should expect people to treat them as discretely quantifiable items (e.g., four objects or four events) and to compare them by means of their cardinality. This possibility goes along with earlier findings that the underlying concepts of entities determine, in part, the dimensions along which people compare them (Barner & Snedeker 2006; Barner et al. 2008). Naturally discrete objects, such as pieces of furniture, promote comparison by number in tasks that ask for explicit judgments of "Who has more furniture?", and naturally punctual events, such as jumps, promote comparison by number in judgments of "Who did more jumps?"

It may also be possible to observe this highlighting of number in situations that don't require explicit quantitative comparisons, and that don't use noun or verb phrases to denote the entities in question—that is, in tasks that may tap more directly the mental representations of the items. Experiment 2 uses a method of this sort that asks participants to rate the similarity of pairs of displays. Pairs that contain naturally atomic items should increase the importance of number in the comparison, relative to pairs that contain non-atomic items.

4. Experiment 2: Non-linguistic similarity

This experiment makes use of participants' judgments of the similarity of two displays to gauge the importance of the number of elements within them. Although there are many psychological theories of similarity, all theories portray similarity as a function of the items' properties and relations. According to some theories (e.g., Tversky 1977), the similarity between two items *x* and *y* depends positively on the number of properties or features that *x* and *y* have in common (and negatively on the number of properties they do not share). According to other theories (e.g., Shepard 1962), similarity depends on the relative positions of *x* and *y* in a dimensionalized psychological space, with greater similarity if *x* and *y* are closer on the relevant dimensions. All theories, however, assume that people's attention to the features or dimensions can shift perceived similarity (Nosofsky 1984; Shepard 1964). For example, consider two objects that have very different shape but relatively similar color, such as a blue circle and a blue-green octagon. A person who attends more to color than to shape will judge these objects more similar to each other than will a person who attends more to shape than to color.

Participants in the present experiment view a pair of displays on each trial, and they rate the similarity of these displays. The members of a pair differ in the number of static or dynamic entities, with the difference in number ranging from one to five. On one trial, for example, a participant might see two displays of petal-shaped forms like those in Figure 6, one display containing four petals and the other containing five. On another trial, a participant might see a pair of displays with a moving entity, like those of Experiment 1, one including four temporal pauses and the other five. In line with the psychophysical models just mentioned, we assume that the difference in the number of discrete forms or movements will have a greater impact on the similarity ratings if participants' attention is drawn to number. If naturally-divided items make number more salient than do unnaturally divided ones, we should find a greater impact of number on the perceived similarity of the former than of the latter.

Thus, our Experiment 2 was a similarity judgment study designed to test how people perceive pairs of images or animations that differ only in their number of divisions and the 'naturalness' of those divisions. The previous experimental research that we have reviewed found that people strongly preferred to label naturally-divided images and animations using plural count nouns (e.g., *some gleebs*) but preferred mass labels for unnaturally-divided images and animations (e.g., *some gleeb/gleebing*). The results of our Experiment 1 with novel verbs reflected the same distinction. Our expectation for Experiment 2 was that the similarity judgments would broadly pattern with the preference results. We expected that perceived similarity of the naturally-divided items should be greater than that for the unnaturally-divided items, due to the greater variety of shapes and paths among the latter. More important, the number of elements in a pair of naturally-divided animations or images should have more impact on similarity than the number of elements in a pair of unnaturally divided items. For example, number should play a bigger part in similarity judgments of the two naturally-divided items at the left of Figure 1 than in judgments of the two unnaturally-divided items at the right. This prediction implies that a plot of the obtained similarity ratings against the difference in number will display a larger slope for the naturally-divided than for the unnaturally-divided stimuli.

[Insert Figure 6 around here]

Participants. We recruited 45 participants through the Northwestern University Department of Psychology subject pool. They received 1 lab credit for 30 minutes of participation. Two participants were excluded: 1 for failure to complete the whole experiment, and 1 due to a computer error. We report the results of the remaining 43 participants.

Method. The method is represented schematically in Figure 6. Participants would first see an image or an animation with n natural or unnatural divisions presented on the screen, followed by another image or animation with m, $n \neq m$, divisions. (Figure 6 represents the case of four versus five naturally-divided items.) Immediately following this, a screen would appear with the question, "How similar were those two images/animations?" (depending on the block), along with a visual representation of the scale to remind participants of its orientation. Participants indicated their judgment by button press using the number keys at the top of the keyboard.

Design. We manipulated the factors DOMAIN (animations, images) and DIVISION (natural, unnatural). In addition, the experiment varied the difference between the number of elements in a pair. We call this difference "CONTRAST," and it ranged from 1 (e.g., a pair consisting of four vs. five objects or movements) to 5 (a pair consisting of four vs. nine objects or movements). We presented our stimuli in a design blocked by the factors DOMAIN and DIVISION, in counter-balanced order. That is, in our design, participants only ever judged (i) pairs of images or pairs of animations, and (ii) pairs of naturally- or unnaturally-divided displays; they were never asked to judge, e.g., an image paired with an animation, or a naturally-divided display paired with an unnaturally-divided display. Participants were furthermore never asked to rate the similarity of identical displays (i.e., there was always a difference in CONTRAST for each pair of displays). Participants were asked to rate "how similar" the two images or animations were on a scale from 1 to 7, with 1 labeled 'not at all similar' and 7 labeled 'very similar'.

Materials. We used the same image and animation stimuli as Experiment 1 in Wellwood et al. (2016), shown schematically in Figures 1 and 2 above. These consisted of items with 4-9 divisions, for each combination of the 2 levels of DOMAIN and DIVISION, for a total of 12 unique images and 12 animations. Each image was visible for 3 seconds, the same amount of time that each animation took on average.

[Insert Figure 7 around here]

Results and discussion. We found that pairs of displays with natural divisions were rated more similar overall, suggesting that participants detected the difference between natural and unnatural divisions, as we predicted. Figure 7 shows that this difference was present for both images and animations. Furthermore, the naturally-divided conditions showed greater sensitivity to numerical contrast—i.e., greater differences in number were seen as more different—as revealed by an interaction effect between DIVISION and CONTRAST. This can be seen as a difference in slopes of the functions in Figure 8, which are the best-fitting straight lines to the average ratings (points in the graph).

We conducted statistical analyses just as in Experiment 1, except we included the factors DOMAIN, DIVISION, and CONTRAST (as a continuous variable) as predictors, and our regressions were linear rather than logistic. As Figure 7 suggests, participants regarded naturally-divided displays as more similar than unnaturally-divided ones. The mean similarity rating for the natural items on our 1-7 scale was 4.36 and the mean for the unnatural items 3.62, $\chi^2 = 49.6$, $p < .001$. As we mentioned earlier, this is probably the result of the uniformity in the shape of the naturally-divided objects and events, compared to the variability in the shape of the unnaturally-divided ones (see Figure 1). However, as Figure 7 also shows, the difference due to naturalness was greater for the images than for the animations, $\chi^2 = 38.8$, $p < .001$. Because an animation unfolds in time, whereas the parts of an image appear simultaneously, participants may have found it easier to perceive variations in similarity in the case of images.

[Insert Figure 8 around here]

The crucial results in this experiment, however, are those associated with the number of items, which appear in Figure 8. Rated similarity decreased across the board as the difference in the number of items increased, $\chi^2 = 95.4$, $p < .001$. This effect is apparent in the negative slopes in the figure. The number effect was also about equally large for animations and for images overall: There was no significant interaction between these two factors, $\chi^2 < 1$. Of more interest, the effect of number was greater for the naturally-divided than for the unnaturally-divided items, $\chi^2 = 11.1$, $p < .001$. For naturally-divided items, similarity ratings decreased by 2.43 scale points from pairs that differed by one element to pairs that differed by five; but for the unnaturally-divided items, similarity ratings decreased by 1.96 scale points across the same range. This lends support to the idea that number is an especially salient dimension of comparison for stimuli that people conceive as atomic, countable units. On our interpretation, the uniform petal-shaped objects or movements drew participants' attention to the number of these items as a main determinant of similarity.

The effect of number on naturally- versus unnaturally-divided items was somewhat greater in the case of images than animations. This appears in Figure 8 as a larger difference in slope between naturally-divided and unnaturally-divided items for the images (at the right of figure) than for the animations (at the left). This difference produced a marginally significant 3-way interaction among CONTRAST (1-5), DOMAIN, (image vs. animation), and DIVISION (natural vs. unnatural), $\chi^2 = 3.5$, $p = .06$. Probing this result further, we found that the effect was driven by an interaction between CONTRAST and DIVISION in the image condition, $\chi^2 = 16.1$, $p < .001$, while there was no such interaction in the animation condition, $\chi^2 = .95$, $p = .33$. We did not predict this pattern, but it may again be due to the possibility, mentioned earlier, that number is more obvious in a simultaneously presented image than in a dynamically unfolding animation. Animations may make more demands on working memory to retain properties like path shape and number, damping the effects of differences among them.

We see these results as confirming the idea that natural divisions make both static and dynamic entities appear atomic and so eligible for counting. The uniform divisions indicate that the entities are distinct, non-overlapping elements to which our standard counting procedures easily apply. The salience of the cardinality of the elements, in turn, makes cardinality apt for comparison. Participants' similarity judgments reflect this effect, since these judgments track cardinality more closely for the naturally-divided displays than for the unnaturally-divided ones. The results dovetail in this way with those of Experiment 1 and the earlier research we cited. Those results indicated that natural divisions prompt count syntax choices; the present results suggest that they do so because the natural divisions render the entities countable.

We note, however, that although our earlier results provide robust evidence that naturally-divided events attract both count nouns (Wellwood et al. 2016) and distributive modifiers (Experiment 1 of the present paper), the effects of natural divisions on similarity in the present experiment is weaker for animations than for images. We've suggested that the weaker effect may be due to the events' greater demands on cognitive processing, but this raises the question of why the same difference did not appear in the earlier studies. Although we have no definitive answer to this question, we suspect the similarity judgments of the present study may be less sensitive than the linguistic judgments of the earlier ones. Similarity judgments, by their nature, are diffuse and flexible (Sloman & Rips 1998): They depend on a particular analysis of the features or dimensions of the to-be-judged items, as we mentioned at the beginning of this section. Variability in the analysis from person-to-person or from time-to-time can limit the ability of such judgments to detect effects of interest.

5. General discussion

We've been assuming that atomistic concepts carve out countable elements across ontological domains. Items that people conceive as having natural or nonabitrary shape present themselves as units that we can count and quantify distributively, no matter whether they are spatial or temporal entities. In the present experiments, we manipulated atomicity through repetition of a single shape that appeared as a bounded region or path, and we found parallel effects in the object and event domains. In Experiment 1, when participants saw animations of an object pausing at natural positions along identically shaped, though invisible paths (i.e., the center of a "flower"), they preferred to describe the event as *gleebing every second or so* in preference to *gleebing around a little*. This result extends to telic (versus atelic) descriptions our earlier findings on count (versus mass) syntax (Wellwood et al. 2016): Repetition of a natural region or path prompts choice of count syntax for both objects and events. Experiment 2 showed that the effects of atomicity extend to a task that does not require overt linguistic judgments. Displaying single-shaped regions or paths highlights the cardinality of these elements and makes differences in cardinality a more important determinant of similarity. Thus, in line with our assumption, people's mental representation of natural divisions in shapes and paths seems to produce an impression of atomic units and make these units individually quantifiable.

The effect of numerical contrast that we observed in our similarity judgment study was perhaps not as large as we might have expected, in light of the hypothesis that the perception of individual units influences the salience of numerical differences. However, one feature of our design may have made this effect more difficult to observe. In both our image and animation conditions, an increase in numerical difference (i.e., the difference between the number of breaks in one image/animation and another image/animation) also represented an increase along other non-numerical dimensions. In the image condition, an increase from four breaks to five breaks was perfectly correlated with an increase in total line length and total pixel coverage between the line drawings. In the animation condition, an increase from four to five breaks was correlated with an increase in total path length and total duration. In light of the earlier experimental results we reviewed above, one possibility is that participants were tuning into the numerical dimension in our natural condition, while tuning into increases along these other dimensions in the unnatural condition. If so, then we might have observed an overall decrease in similarity scores as numerical contrast increased, but for different reasons in the two conditions. We are exploring this possibility in ongoing work by comparing conditions in which the numerical dimension is correlated and anti-correlated with continuous dimensions like length (images) and duration (animations).

A further suggestion for future work raised by these studies is the possibility of using the properties of conceptualization—in terms of atomic objects and events, as opposed to non-atomic substances and processes—to predict how people should quantify with novel nouns and verbs. Barner and Snedeker (2006) showed that adults and children can make use of conceptual features of static entities as well as their knowledge of the morphosyntactic mass/count distinction to influence their judgments of what counts as "more" with a novel noun like *fem*. In one condition, for example, they would present adults and children with a portion of a non-solid substance for which the participants likely did not already have a name (e.g., green butter). It would be described using a noun phrase like *a fem* (count syntax) or *some fem* (mass syntax). They found that adults categorically preferred to judge presentations of multiple such portions using *more fems* based on number, but *more fem* based on volume. The three year olds showed a similar asymmetry, but showed a stronger bias towards volume. Our results suggest the possibility that parallel preferences could obtain for dynamic entities, whether described using deverbal nouns (e.g., *do more femming*/*do more fems*; cf. Barner, Wagner & Snedeker 2008) or a novel verb (e.g., *gleeb more*). Given dynamic displays with (what we have called) unnatural divisions, we would expect count syntax to unambiguously bias participants towards quantification by the number of breaks, but we would expect participants to be biased towards continuous dimensions when the comparison is expressed using mass syntax or using a novel verb (unmarked for number).

These results also have implications for the early cognitive development, prior to successful acquisition of the relevant grammatical distinctions. Our contention is that the conceptual categories that we have targeted are not derivative of linguistic knowledge, but rather present and available independently. A wealth of evidence from developmental science suggests that objects, substances, events, and processes are fundamental to how we shape our experience of the world, and are evident to some degree from the earliest ages it is possible to test (Ferry et al. 2015; Hespos et al. 2009a; 2009b; 2010). Results from preverbal infants provide the crucial sort of evidence for any claim of conceptual priority, since in these cases it is not obviously possible to attribute the knowledge to some abstraction based on linguistic knowledge. Rather, language learning seems to develop by linking linguistic forms to universal, pre-existing conceptual categories (Hespos & Spelke 2004). We are currently exploring whether preverbal infants reveal similar patterns to adults when tested with images and animations that vary in precisely the ways we have tested in the experiments reported here.

This paper thus attempts to contribute to recent research suggesting a “naturalization” of natural language semantics. This project aims to ground at least some of the semanticists’ formal posits in independently-understood aspects of human psychology. To the extent that we have been successful, the project hints at a view wherein the categories of entity posited in our models—the objects, substances, events, processes, states, and other things besides—correspond to categories of conceptualization whose properties, in many cases, have yet to be isolated in cognition. If the properties that semanticists posit for such entities predict properties of conceptualization, independently of language (potentially even before a language has been acquired, as in the case of prelinguistic infants), the question of what sort of meaning theory predicts such generalizations becomes more acute. At a minimum, establishing these generalizations could suggest a view on which semantic interpretation in Lewis’ (1970) sense—as establishing a direct linkage between an expression and the mind-independent world—is seriously attenuated. The world might not deliver the right sorts of entities for that interpretation.

References

Acquaviva, Paolo. 2008. *Lexical plurals: A morphosemantic approach*. Oxford, UK: Oxford University Press.

Bach, Emmon. 1986a. Natural language metaphysics. In Ruth Barcan Marcus, Gorg J. W. Dorn, and Paul Weingartner (eds.) *Logic, methodology and philosophy of science VII*. Amsterdam: North-Holland. 573–595.

Bach, Emmon. 1986b. The algebra of events. *Linguistics and Philosophy* 9(1). 5–16.

Bach, Emmon and Wynn Chao. 2012. The metaphysics of natural language(s). In Ruth Kempson, Tim Fernando and Nicholas Asher (eds.) *Handbook of philosophy of science. Volume 14: Philosophy of linguistics*. London: Elsevier. 175–196.

Barner, David and Jesse Snedeker. 2006. Children’s early understanding of mass-count syntax: Individuation, lexical content, and the number asymmetry hypothesis. *Language Learning and Development* 2(3). 163–194.

Barner, David, Laura Wagner and Jesse Snedeker. 2008. Events and the ontology of individuals: Verbs as a source of individuating mass and count nouns. *Cognition* 106. 805–832.

Barr, Dale J., Roger Levy, Christoph Scheepers and Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68. 255–278.

Boolos, George. 1981. For every A there is a B. *Linguistic Inquiry* 12(3). 465–467.

Borer, Hagit. 2005. *In name only.* Oxford, UK: Oxford University Press.

Casati, Roberto and Achille C. Varzi. 2008. Event concepts. In Thomas Shipley and Jeffrey Zacks (eds.) *Understanding events* (ch. 2). New York: Oxford University Press. 31–53.

Champollion, Lucas. 2017. *Parts of a whole: Distributivity as a bridge between aspect and measurement*. Oxford: Oxford University Press.

Champollion, Lucas and Manfred Krifka. 2016. Mereology. In Maria Aloni and Paul Dekker (eds.) *The Cambridge handbook of formal semantics.* Cambridge, UK: Cambridge University Press. 512-541.

Chierchia, Gennaro. 1998. Plurality of mass nouns and the notion of "semantic parameter". In Susan Rothstein (ed.) *Events and grammar*. Norwell, MA: Kluwer Academic Publishers. 53–103.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Massachusetts: MIT Press.

Chomsky, Noam. 1995. Language and nature. *Mind* 104(403). 1–61.

Davidson, Donald. 1967. The logical form of action sentences. In Nicholas Rescher (ed.) *The logic of decision and action*. Pittsburgh: Pittsburgh University Press. 81–95.

Davidson, Donald. 1985. Adverbs of action. In Bruce Vermazen and Merrill B. Hintikka (eds.) *Essays on Davidson: Actions and events*. Oxford: Clarendon Press. 230-241

Ferry, Alissa, Susan J. Hespos and Dedre Gentner. 2015. Prelinguistic relational concepts: Investigating analogical processing in infants. *Child Development* 86(5). 1386–1405.

Gillon, Brendan S. 1992. Towards a common semantics for English count and mass nouns. *Linguistics and Philosophy* 15(6). 597–639.

Gillon, Brendan S. 2012. Mass terms. *Philosophy Compass* 7(10). 712–730.

Grimm, Scott. 2012. *Number and individuation*. Doctoral dissertation. Stanford University.

Henderson, Robert. 2013. *Ways of pluralizing events*. Doctoral dissertation. University of California, Santa Cruz.

Hespos, Susan J. and Elizabeth S. Spelke. 2004. Conceptual precursors to spatial language. *Nature* 430. 453–456.

Hespos, Susan J., Alissa Ferry and Lance Rips. 2009a. Five-month-old infants have different expectations for solids and liquids. *Psychological Science* 20(5). 603–611.

Hespos, Susan J., Megan M. Saylor and Stacy Grossman. 2009b. Infants' ability to parse continuous actions. *Developmental Psychology* 45(2). 575–585.

Hespos, Susan J., Stacy Grossman and Megan M. Saylor. 2010. Infants' ability to parse continuous action series: Further evidence. *Neural Networks* 23. 1026–1032.

Hoepelman, Jakob and Christian Rohrer. 1980. On the mass count distinction and the French imparfait and passé simple. In Christian Rohrer (ed.) *Time, tense and aspect*. Tübingen: Niemeyer. 629–645.

Koslicki, Katherin. 1997. Isolation and non-arbitrary division: Frege's two criteria for counting. *Synthese* 112(3). 403–430.

Krifka, Manfred. 1989. Nominal reference, temporal constitution and quantification in event semantics. In Renate Bartsch, Johan van Benthem and Peter van Emde Boas (eds.) *Semantics and contextual expression*. Dordrecht: Foris. 75–115.

Lewis, David. 1970. General semantics. *Synthese* 22. 18–67.

Lidz, Jeffrey, Justin Halberda, Paul Pietroski and Tim Hunter. 2011. Interface transparency and the psychosemantics of *most*. *Natural Language Semantics* 6(3). 227–256.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In Rainer Bäuerle, Christoph Schwarze and Arnim von Stechow (eds.) *Meaning, use and interpretation of language*. Berlin: de Gruyter. 302–323.

Mathieu, Éric. 2012. Flavors of division. *Linguistic Inquiry* 43(4). 650–679.

Montague, Richard. 1973. The proper treatment of quantification in ordinary English. In Richmond H. Thomason (ed.) *Formal philosophy: Selected papers of Richard Montague.* New Haven, Connecticut: Yale University Press. 247–270.

Mourelatos, Alexander. 1978. Events, processes, and states. *Linguistics and Philosophy* 2. 415–434.

Nosofsky, Robert M. 1984. Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 10. 104–114.

Ojeda, Almerindo. 2005. The paradox of mass plurals. In Salikoko S. Mufwene, Elaine J. Francis and Rebecca Wheeler (eds.) *Polymorphous linguistics*: *Jim McCawley's legacy*. Cambridge, MA: MIT Press. 389–410.

Partee, Barbara. 1979. Semantics – mathematics or psychology? In Rainer Bäuerle, Urs Egli and Arnim von Stechow (eds.) *Semantics from different points of view*. Berlin: Springer. 1–14.

Pelletier, Francis J. 2011. Descriptive metaphysics, natural language metaphysics, Sapir-Whorf, and all that stuff: Evidence from the mass-count distinction. *The Baltic International Yearbook of Cognition, Logic and Communication*: Vol. 6. 1–46. https://doi.org/10.4148/biyclc.v6i0.1570

Pietroski, Paul. 2010. Concepts, meanings, and truth: First nature, second nature, and hard work. *Mind & Language* 25(3). 247–278.

Pietroski, Paul. 2015. Framing event variables. *Erkenntnis* 80(1). 31–60.

Pietroski, Paul, Jeffrey Lidz, Tim Hunter and Justin Halberda. 2009. The meaning of *most*: semantics, numerosity, and psychology. *Mind & Language* 24. 554–585.

Prasada, Sandeep, Krag Ferenz and Todd Haskell. 2002. Conceiving of entities as objects and as stuff. *Cognition* 83. 141–165.

Rips, Lance and Susan J. Hespos. 2015. Divisions of the physical world: Concepts of objects and substances. *Psychological Bulletin* 141(4). 786–811.

Rothstein, Susan. 1995. Adverbial quantification over events. *Natural Language Semantics* 3. 1–31.

Rothstein, Susan. 2004. *Structuring events*. Oxford: Blackwell Publishing.

Rothstein, Susan. 2010. Counting and the mass/count distinction. *Journal of Semantics* 27. 343–397.

Rothstein, Susan. 2017. *Semantics for counting and measuring*. Cambridge, UK: Cambridge University Press.

Shepard, Roger N. 1962. The analysis of proximities: Multidimensional scaling with an unknown distance function. Part I. *Psychometrika* 27. 125–140.

Shepard, Roger N. 1964. Attention and the metric structure of stimulus space. *Journal of Mathematical Psychology* 1. 54–87.

Sloman, Steven A. and Lance Rips. 1998. Similarity as an explanatory concept. *Cognition* 65. 87–101.

Tversky, Amos. 1977. Features of similarity. *Psychological Review* 84. 327–352.

Vendler, Zeno. 1957. Verbs and times. *Philosophical Review* 66. 143–160.

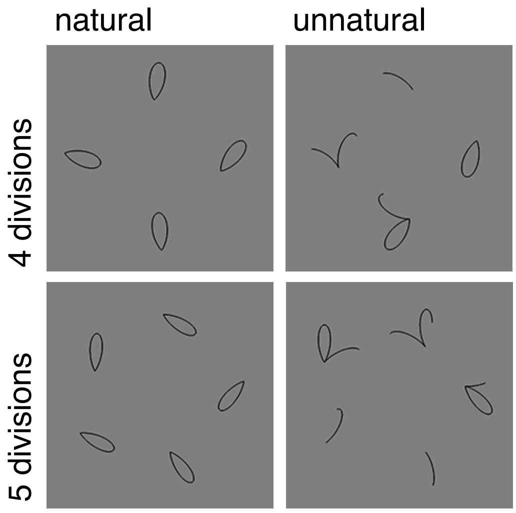
Vogel, Christopher A., Alexis Wellwood, Rachel Dudley and J. Brendan Ritchie. 2014. Talking about causing events. *The Baltic International Yearbook of Cognition, Logic, and Communication*: Vol. 9. 1–22. https://doi.org/10.4148/1944-3676.1092

Wellwood, Alexis. 2014. *Measuring predicates*. Doctoral dissertation. University of Maryland, College Park.

Wellwood, Alexis, Susan J. Hespos and Lance Rips. 2016. The *object : substance :: event : process* analogy. To appear in Tania Lombrozo, Joshua Knobe and Shaun Nichols (eds.) *Oxford studies in experimental philosophy* (vol. 2). Oxford, UK: Oxford University Press.

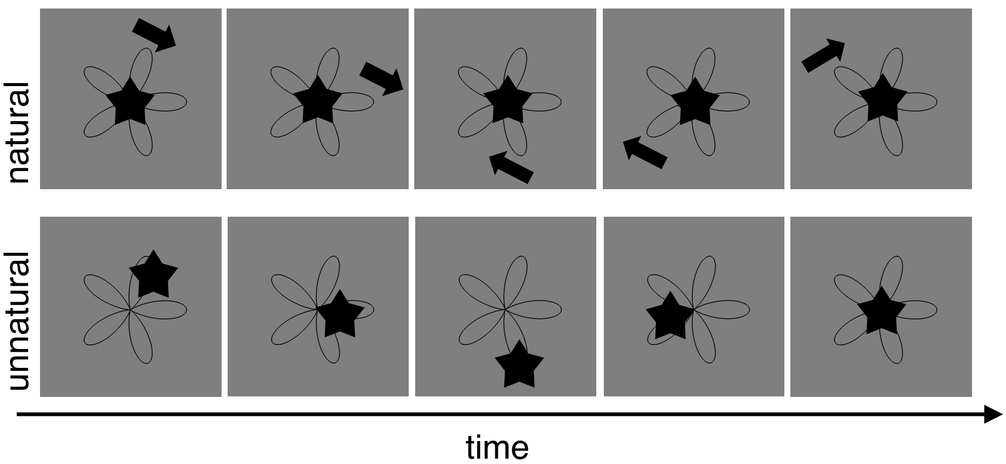
Zacks, Jeff M. and Barbara Tversky. 2001. Event structure in perception and conception. *Psychological Bulletin* 127. 3–21.

Zucchi, Sandro and Michael White. 2001. Twigs, sequences and the temporal constitution of predicates. *Linguistics and Philosophy* 24. 223–270.



change the gray
of the rectangles
(too dark)

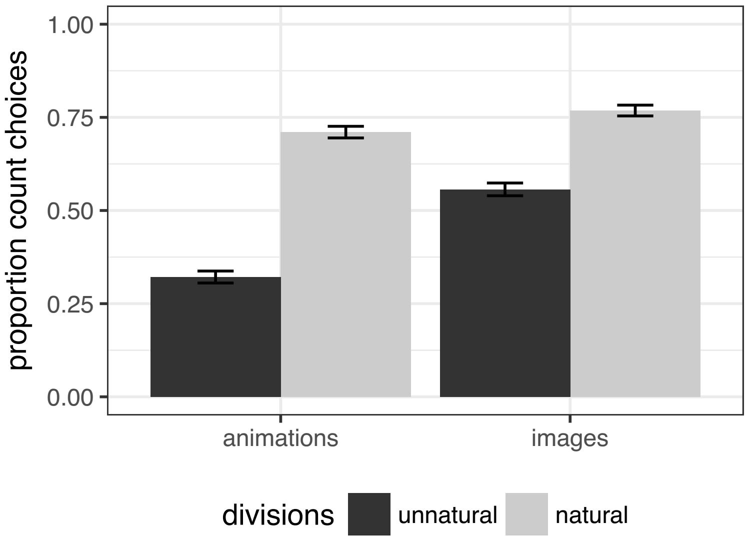
Figure 1. Samples of naturally-divided images with 4 and 5 breaks (left) and unnaturally-divided images with 4 and 5 breaks (right).



change the gray
of the rectangles
(too dark)

Figure 2. Schematic of a naturally-divided animation with 5 breaks (top) and an unnaturally-

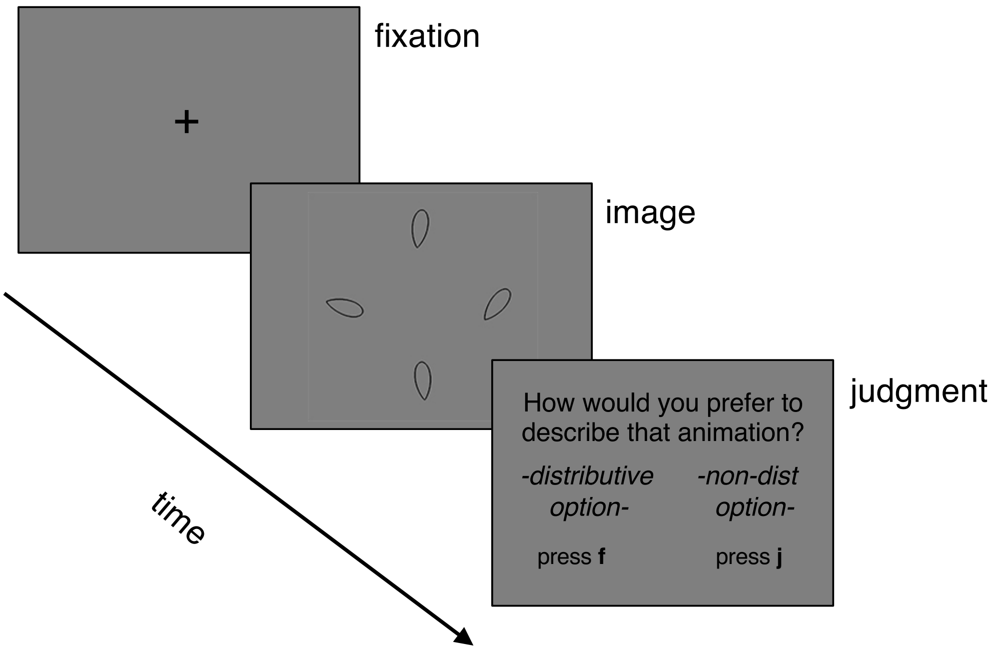
divided animation with 5 breaks (bottom). The position of the star in each panel represents a point where it pauses as it traverses the flower-shaped path (the path itself was invisible to participants).



change the bar chart colours from black to gray, and the gray to white (with a black frame)

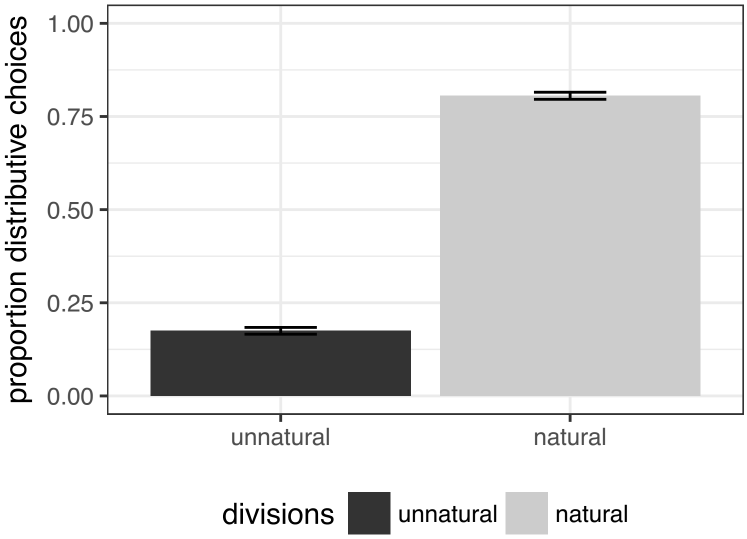
telic2-graph-DomTel-2.pdf

Figure 3. Results from Wellwood et al. (2016). Overall, count syntax was chosen more for images than animations, and for naturally-divided displays than for unnaturally-divided displays. There were no interaction effects, suggesting that the naturalness of divisions influenced categorization in the same way across domains. Error bars represent the standard error of the mean.



change the gray of the rectangles (too dark)

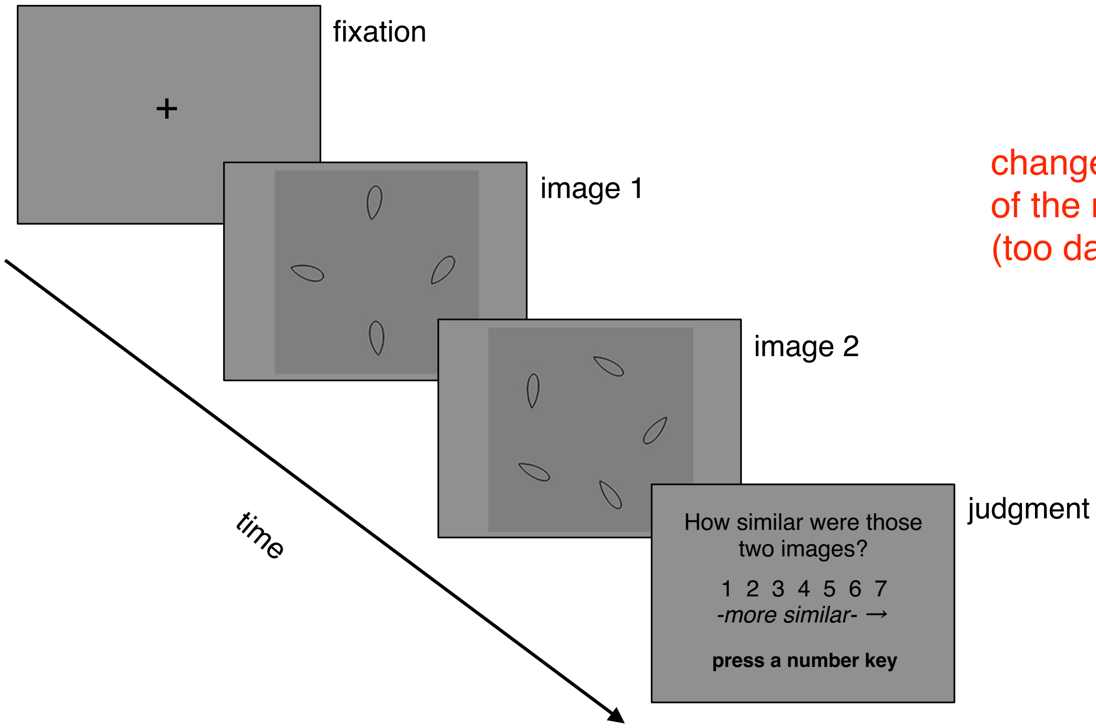
Figure 4. Trial structure for Experiment 1.



change the bar chart colours from black to gray, and the gray to white (with a black frame)

telicY-graph-Tel-2.pdf

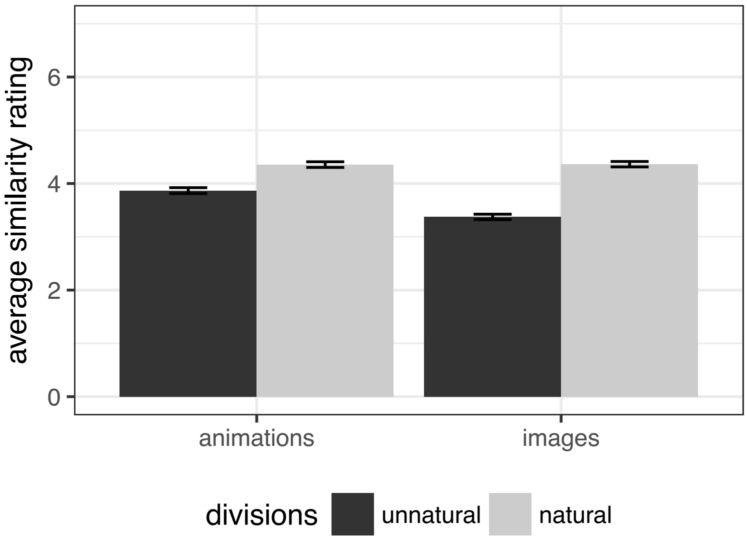
Figure 5: Experiment 1 results. Participants more strongly preferred to pair the sentence with *gleeb every second or so* with the naturally-divided animations than with the unnaturally-divided animations. Error bars represent the standard error of the mean.



change the gray of the rectangles (too dark)

Figure 6. Schematic of the trial structure for Experiment 2, comparing 4- and 5-naturally

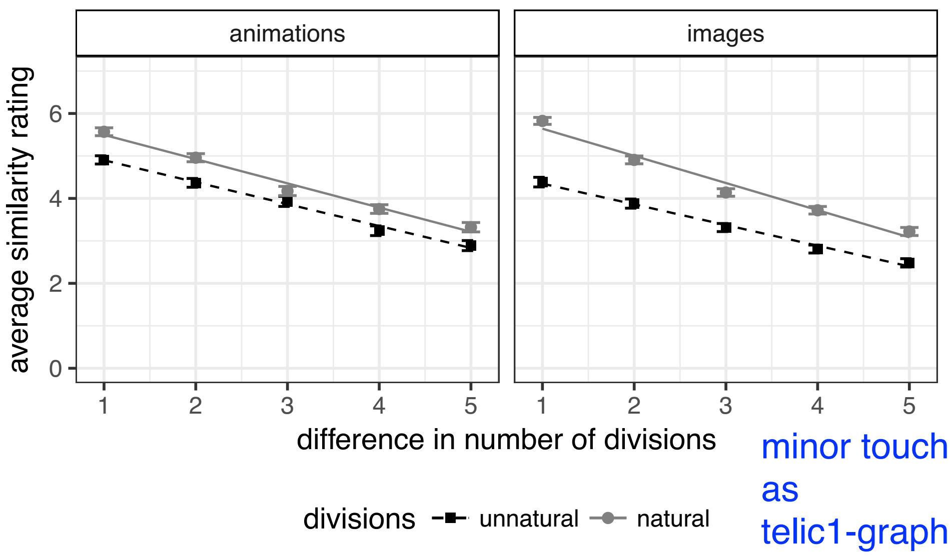
divided images.



change the bar chart colours from black to gray, and the gray to white (with a black frame)

telic1-graph-DomTel-2.pdf

Figure 7. Experiment 2 results: similarity by domain and divisions. Overall, pairs of animations were judged more similar than pairs of images, and the naturally-divided animations and images were judged more similar than the unnaturally-divided animations and images. Also, the difference between natural and unnatural was greater for images than for animations. Error bars represent the standard error of the mean.



minor touch-ups as telic1-graph-DomTelLoop-2.pdf

Figure 8. Experiment 2 results: similarity by domain, divisions, and contrast in number. The

same effects as revealed in Figure 7 are evident here, as is the greater difference in slopes for the naturally- divided animations and images than for the unnaturally-divided animations and images.