LANGUAGE AND COGNITIVE PROCESSES, 2012, iFirst, 1–35

Psychology Press
Taylor & Francis Group

Verb argument structure predicts implicit causality: The advantages of finer-grained semantics

Joshua K. Hartshorne and Jesse Snedeker

Department of Psychology, Harvard University, Cambridge, MA, USA

While the referent of a nonreflexive pronoun clearly depends on context, the nature of these contextual restrictions is controversial. The present study seeks to characterise one representation that guides pronoun resolution. Our focus is an effect known as "implicit causality". In causal dependant clauses, the preferred referent of a pronoun varies systematically with the verb in the main clause (contrast *Sally frightened Mary because she*... with *Sally feared Mary because she*...). A number of researchers have tried to explain and predict such biases with reference to semantic classes of verbs. However, such studies have focused on a small number of specially selected verbs. In Experiment 1, we find that existing taxonomies perform near chance at predicting pronoun-resolution bias on a large set of representative verbs. However, a more fine-grained taxonomy recently proposed in the linguistics literature does significantly better. In Experiment 2, we tested all 264 verbs in two of the narrowly defined verb classes from this new taxonomy, finding that pronoun-resolution biases were categorically different. These findings suggest that the semantic structure of verbs tightly constrains the interpretation of pronouns in causal sentences, raising challenges for theories which posit that implicit causality biases reflect world knowledge or arbitrary lexical features.

Keywords: Pronoun resolution; Implicit causality; Thematic roles; Psychological predicates; Psych verbs; Predicate decomposition.

A proper name like *Catherine the Great* almost always refers to the same person: Catherine the Great. In contrast, a third-person pronoun like *she* can refer to a different entity each time it is used; thus the referent must be fixed by information in the context in which the pronoun is used. Some contextual cues—like pointing to an individual while uttering the pronoun (Nappa & Arnold, 2009)—simply pick out the referent by directing the listener's attention to particular entity. The representational basis of other contextual cues, however, is less obvious. For instance, most English-speakers resolve the pronoun to Sally in (1) but to Mary in (2):

Correspondence should be addressed to Joshua Hartshorne, Department of Psychology, Harvard University, 33 Kirkland Street, WJH 1120, Cambridge, MA 02138, USA. E-mail: jharts@wjh.harvard.edu

The authors wish to thank Alfonso Caramazza, Susan Carey, Steve Pinker, Manizeh Khan, Mahesh Srinivasan, Nathan Winkler-Rhoades, Rebecca Nappa, and three anonymous reviewers for comments and discussion. This material is based on work supported by a National Defense Science and Engineering Graduate Fellowship to JH and a grant from the National Science Foundation to JS (0623845).

http://www.psypress.com/lcp http://dx.doi.org/10.1080/01690965.2012.689305

(1) Sally frightens Mary because she is a strange girl.
(2) Sally fears Mary because she is a strange girl.

This contrast suggests that the verb in the main clause can influence the resolution of the pronoun in the second clause, with "subject-biased" verbs like *frighten* leading to subject interpretations, while "object-biased" verbs like *fear* lead to object interpretations. This systematic difference between verbs cannot be solely attributed to the plausibility of the material after the pronoun. These verb effects are apparent even in cases where the sentence is cut off after the pronoun or the second clause has no meaningful content:

(3) Sally frightens Mary because she is a dax.
(4) Sally fears Mary because she is a dax.

This is a bias, however, and not an absolute constraint on our interpretation of the pronoun. For example, (5) and (6) are both coherent and plausible, though inconsistent with the bias, indicating that the content of the second clause can override the bias introduced by the verb:

(5) Sally frightens Mary because she [Mary] is so timid.
(6) Sally fears Mary because she [Sally] is so timid.

This systematic shift in pronoun interpretation depending on the verb is accompanied by a shift in our interpretation of the causal structure of the sentences. The notion of causation appears to be crucial. If the two clauses are linked by some other discourse relation (e.g. the second is a consequent rather than a cause) then the pronoun resolution biases will differ (Crinean & Garnham, 2006; Ehrlich, 1980; Kehler, Kertz, Rodhe, & Elman, 2008; Stewart, Pickering, & Sanford, 1998).

In a related finding, Brown and Fish (1983a) reported that people believe that Sally is responsible for Mary's fright in sentences like (7), and that Mary is responsible for Sally's fear in sentences like (8):

(7) Sally frightens Mary.
(8) Sally fears Mary.

Many of the verbs that have been found to be subject or object biased in pronoun resolution tasks have also been found to be subject or object biased, respectively, in the Brown and Fish causal attribution task (Rudolph & Forsterling, 1997). For this reason, both these effects, and a host of closely related phenomena, have been treated as a single construct which is generally called "implicit causality" (Brown & Fish, 1983a; Garvey & Caramazza, 1974; Garvey, Caramazza, & Yates, 1974).

Implicit causality (IC) provides a potential link between a linguistic process (pronoun resolution) on the one hand, and our conceptual understanding of the causal structure of the world and social relationships on the other. For this reason, it is of interest to a wide range of researchers. The phenomenon has been used to probe the development of causal schemas in children (Au, 1986; Corrigan & Stevenson, 1994), the stability of these schemas across cultures (Brown & Fish, 1983b), and the conceptualisation of social relationships and dominance hierarchies (Corrigan, 2001; LaFrance, Brownell, & Hahn, 1997; Maass, Salvi, Arcuri, & Semin, 1989; Mannetti & De Grada, 1991). Psycholinguists have used IC as a test case for studying the interplay

of bottom-up and top-down processing in language comprehension (Featherstone & Sturt, 2010; Garnham, Traxler, Oakhill, & Gernsbacher, 1996; Greene & McKoon, 1995; Guerry, Gimenes, Caplan, & Rigalleau, 2006; Koornneef & Van Berkum, 2006; Long & De Ley, 2000; McDonald & MacWhinney, 1995; McKoon, Greene, & Ratcliff, 1993; Pyykkonen & Jarvikivi, 2010; Shen & Yang, 2006; Stewart, Pickering, & Sanford, 2000) and the developmental origins of these processes (Pyykkonen, Matthews, & Jarvikivi, 2010). While some of these researchers have approached IC as an isolated phenomenon, others have addressed it as part of a broader theory of discourse coherence, treating it as a specific example of how the interpretation of one sentence is constrained by its relation to other sentences in the discourse (Crinean & Garnham, 2006; Ehrlich, 1980; Frank, Koppen, Noordman, & Vonk, 2007; Kehler et al., 2008; Pickering & Majid, 2007; Stewart et al., 1998). Other researchers have asked whether IC is an effect of language on thought or of thought on language (Brown & Fish, 1983a; Hoffman & Tchir, 1990) *a la* the Sapir–Whorf hypothesis (Whorf, 1956). Yet others have interpreted IC as an example of a cognitive heuristic, or short cut, and have used it to explore the effect of mood on the use of heuristics (De Goede et al., 2009; cf. Forgas, 1995).

In the remainder of the introduction we review previous theoretical accounts of the mechanisms underlying IC, discuss reasons for revisiting one of these accounts (the *semantic structure* account), and provide an overview of the present experiments.

WHAT IS IMPLICIT CAUSALITY?

Despite its wide application in psychology, the nature of IC is very poorly understood. While most authors agree that a verb's bias is due to its meaning, they disagree as to what aspect of meaning is relevant (Au, 1986; Brown & Fish, 1983a; Corrigan, 2001; Crinean & Garnham, 2006; Garvey & Caramazza, 1974; Garvey et al., 1974; LaFrance et al., 1997; Maass et al., 1989; Mannetti & De Grada, 1991; Rudolph & Forsterling, 1997; Semin & Fiedler, 1988, 1991). While there are many differences between proposals, accounts can be roughly grouped into three types: the *arbitrary semantic tag* account, the *world knowledge* account, and the *semantic structure* account.

On the *arbitrary semantic tag* account, verbs have an implicit cause feature which does not reduce to any other aspect of the verb (e.g. it is not predicted by independent semantic, syntactic, or other features of/facts about the verb). Rather, each verb is marked in its lexical entry as being either subject biased or object biased, just as nouns in many languages are marked for gender. This serves as the null hypothesis about the basis of IC biases (that there is no other factor which predicts them) and is most explicit in work by Garvey and Caramazza (Caramazza, Grober, Garvey & Yates, 1977; Garvey et al., 1974).

On the *world knowledge account*, the implicit cause of the verb is inferred from learned distributional facts about situations to which the verb is typically applied. Much of this work is influenced by Brown and Fish's (1983a) demonstration that IC pronoun biases correlate with offline judgments of causal responsibility (see 7 & 8, above). Subsequent studies also demonstrated that the IC bias is modulated by the relative social status of the characters in the sentences, with higher status individuals being seen as more causal (Corrigan, 2001; LaFrance et al., 1997; Mannetti & De Grada, 1991). These observations led many researchers to propose that IC is an inference based on nonlinguistic knowledge about who tends to cause certain types of

events, rather than being derived directly from the literal meaning of the verb (e.g. Au, 1986; Koornneef & Van Berkum, 2006; Maass et al., 1989). That is, verbs do not encode any information about causality; rather, causality is inferred.

In contrast, the *semantic structure* account posits that IC bias is derived from the literal meaning of the verb. So while on the *world knowledge* account, changing facts about the world changes the IC bias of a verb, on the *semantic structure* account, facts about the world are relevant primarily in that they cause speakers to coin verbs that encode specific causal information. Brown and Fish (1983a) proposed the first such theory, systematically relating each verb's bias (subject biased, like *frighten*, or object biased, like *fear*) to the thematic roles of the verb's arguments. In linguistic theory, thematic roles are invoked to explain how the different arguments of a verb are syntactically realised and the range of syntactic contexts in which a verb can appear (for review, see Levin & Rappaport Hovav, 2005). Brown and Fish distinguished four thematic roles and a total of three types of verbs: agent-patient ("action") verbs, stimulus-experiencer verbs, and experiencer-stimulus verbs.

On this account, the subject of an action verb (*kick, paint, break, throw*) is an AGENT which effects some change on the PATIENT.[1] Because AGENTS are by definition causal actors (they cause the PATIENT to change), pronouns in sentences like (1–8) should be biased towards the subjects of action verbs. In contrast, there are two distinct patterns for transitive verbs that describe psychological states (called "psych verbs"). For some psych verbs (e.g. *frighten, confuse, surprise*) the subject of the verb is a STIMULUS, which elicits a psychological state in the verb's object (the EXPERIENCER). For other psych verbs, these roles are reversed: the subject is the EXPERIENCER and the object is the STIMULUS (e.g. *like, love, hate*). Like AGENT, the notion of STIMULUS is explicitly causal; therefore Brown and Fish argued that IC bias should follow the STIMULUS. Thus the stimulus-experiencer verbs (*frighten, confuse, surprise*) are subject biased, whereas experiencer-stimulus verbs (*like, love, hate*) are typically object biased.

The primary difficulty with this account is that numerous exceptions to these predictions have been found. Brown and Fish's (1983a) verb taxonomy has been revised several times, with mixed results. For example, Au (1986) identified numerous action verbs such as *thank* and *punish* which are object biased, contra Brown and Fish's predictions. These verbs were further investigated by Rudolph and Forsterling (1997), who labelled them "agent-evocator" verbs but were unable to determine anything that distinguished them from other action verbs beyond the fact that they are object biased. In contrast, Semin and Fiedler (1988, 1991) suggested that subject-biased action verbs are those which do not entail a specific physical action (*stop, help*), whereas verbs which do entail specific physical actions (*punch, kick*) are only weakly (subject-)biased (note that their taxonomy, while similar to those above, does not invoke thematic roles). One obvious challenge for this taxonomy is that it does not predict any object-biased action verbs, despite considerable evidence that such verbs exist (Au, 1986; Rudolph & Forsterling, 1997). In addition, the fact that some researchers have identified verbs of intermediate bias (e.g. Garvey et al., 1974; Semin & Fiedler, 1991) has raised questions about apparently more categorical taxonomic theories.

[1] Throughout, SMALL CAPS are used to refer to thematic roles.

THE PRESENT STUDY

Though there has been relatively little work on IC verb taxonomies in recent years (but see Crinean & Garnham, 2006; Ferstl, Garnham, & Manoulidou, 2011), there are reasons to revisit the claim that semantic structure plays a role in pronoun resolution. First, the twenty-eight years since Brown and Fish (1983a) have witnessed an explosion of research in lexical semantics, with many semanticists arguing for considerably more complex and nuanced accounts of the semantic structure of verbs (Goldberg, 1995; Jackendoff, 1990; Levin, 1993; Pinker, 1989; Pustejovsky, 1995; Talmy, 2000a, 2000b; Van Valin, 2004; see Levin & Rappaport Hovav, 2005, for a review). While previous *semantic structure* accounts were based on relatively simple linguistic theories involving 4–5 thematic roles, current theories invoke a much richer set of theoretical primitives, allowing for more fine-grained distinctions. This can be achieved by positing a larger set of thematic roles (Baker, 1988; Kipper, Korhonen, Ryant, & Palmer, 2008; Pesetsky, 1995), by treating thematic roles as prototypes based on a larger set of semantic features (Dowty, 1991), or by building verb semantics out of combinations of primitive predicates (Harley, 1995; Moens & Steedman, 1988; Pinker, 1989; Van Valin, 1993). Thus, in the analyses below, we explore whether more modern theories of semantics can provide additional insight into IC. Second, most studies of IC bias have focused on a small set of verbs that have been repeatedly sampled (Rudolph & Forsterling, 1997), an approach which may have masked patterns in IC biases. Finally, there is a large body of work suggesting that much of the syntax-semantics interface is organised around notions of causality, and that causal information is encoded in the semantics of verbs in both English and other languages (Ambridge, Pine, Rowland, & Young, 2008; Croft, 1991, in press; Lidz, Gleitman, & Gleitman, 2003; Naigles, 1990; Pesetsky, 1995; Pinker, 1984, 1989; Talmy, 2000a, 2000b). If these analyses are correct, it would be remarkable if such information did not play a role in IC, and, in fact, there is no evidence *against* semantic structure playing a role in IC biases. The claim that nonlinguistic factors such as social dominance also affect pronoun resolution has been taken as evidence in favour of the *world knowledge* account, but data of this kind are not inconsistent with the possibility that semantic structure plays an independent and central role in generating causal inferences.[2]

The present study attempts to address these deficiencies in several ways. First, in order to provide a better understanding of the distribution of IC biases, in Experiment 1 we elicit biases for 720 high-frequency verbs, which were chosen irrespective of causal structure or IC bias. We find that earlier IC verb taxonomies fail to capture the pattern of results. In particular, in each taxonomy there are proposed categories which include a even mix of subject-biased verbs and object-biased verbs. Nonetheless, when we reanalyse these data according to more recent theories of semantics, we find new subclasses of verbs that have consistent IC biases. We then test the reliability and generalisability of these findings.

On the *semantic structure* hypothesis, there should be a relationship between the encoded meaning of a given verb and its IC bias. The fact that there is active debate about the nature of verbal structures (Levin & Rappaport Hovav, 2005) presents a challenge, as an account based on any particular theory could prove wrong even though the general approach is correct. We attempt to circumvent this problem by

[2] Interestingly, social dominance is only known to affect causal attribution tasks, such as those used by Brown and Fish (1983a). There is little evidence that such manipulations affect pronoun resolution (Goikoetxea et al., 2008; Hartshorne, *submitted*; but see Ferstl et al., 2011).

classifying verbs not based on primitives specific to a particular theory, but based on the data that all such theories attempt to predict: the sentence frames a given verb can appear in.[3] In particular, we classify our verbs according to VerbNet, which is an extension of Levin's (1993) analysis of verb argument structure applied to more verbs and supplemented with refinements suggested by Korhonen and Briscoe (2004) and Korhonen and Ryant (2005).[4] A total of 5,879 verbs are divided in 274 classes, primarily distinguished based on the types of arguments the verbs can take. Thus *please* and *like* are in classes 31.1 and 31.2, respectively, because the latter can take a *that* complement (*I liked that he was so honest*) while the former cannot (**I pleased that he was so honest*). Critically, verb classes consist of verbs that appear not just in the same individual frame but in the same set of frames (or alternations). For instance, the verbs in class 31.2 can appear in the following frames:

(9) a. NP V NP (The tourists liked the painting.)
b. NP V NP [for] NP (I liked him for his honesty.)
c. NP V NP [in] NP (I liked the honesty in him.)
d. NP V [that] S (I liked that he was so honest.)
e. NP V S-ing (I like writing.)
f. NP V NP S-ing (I like him wearing suits.)

Finally a small number of verb classes were divided according to what Kipper et al. (2008) call "considerations of meaning". For instance, while *convert to* and *turn to* are highly similar in argument structure, in (10) the straw actually becomes gold, while in (11) the natives do not become Deism:

(10) The fairy turned the straw to gold.
(11) The missionaries converted the natives to Deism.

Importantly, these verb classes are empirical—rather than theoretical—constructs, which must be explained on any theory. Most theories of lexical semantics try to account for the existence of these classes by making reference to a smaller number of features which, in combination, generate all the classes. On some accounts, for instance, the syntactic behaviour of a verb class is a product of the thematic roles the verbs in the class take or the sublexical predicates with which the verbal root can compose (Levin & Rappaport Hovav, 2005). For IC *semantic structure* accounts, what is important are these underlying features, not the verb class *per se*. Thus, the VerbNet classes may well be *more* fine-grained than is necessary. Nonetheless, we focus on these fine-grained verb classes because on all *semantic structure* accounts, verbs within a class should behave uniformly with respect to IC, whereas which verb classes should behave similarly to one another may vary across specific theoretical proposals.

[3] Not all theories of semantics are focused on such issues, and in principle IC bias could be related to an aspect of verbal semantics not relevant to the verb's distribution across sentence frames. We nonetheless focus on these frame-relevant semantic theories for three reasons. First, this is the notion of semantic structure invoked in IC *semantic structure* accounts such as Brown and Fish (1983a, 1983b). Second, the distribution of verbs across sentence frames is an objective, easily observable phenomenon. Third, although theories of verb meaning vary along many dimensions, the same classes of verbs appear repeatedly across different theories (Goldberg, 1995; Jackendoff, 1990; Levin, 1993; Pinker, 1989; Pustejovsky, 1995; Talmy, 2000a, 2000b; Van Valin, 2004).

[4] The primary difference is the addition of nearly 200 subclasses. These subclasses play only a minimal role in the present analyses. In addition, the lists of verbs in each class are more extensive in VerbNet.

However, in the General Discussion, we also evaluate two specific approaches to verb semantics.

Apart from any theoretical considerations, VerbNet offers notable advantages over other approaches. It is by far the most comprehensive semantic classification scheme for verbs and is available in electronic format on the Internet (verbs.colorado.edu/ ~mpalmer/projects/verbnet.html).

EXPERIMENT 1

In Experiment 1, we elicit IC pronoun biases for 720 high-frequency verbs. As shown in Table 1, existing IC verb taxonomies predict that most (Semin & Fiedler, 1991) or all (Brown & Fish, 1983a; Rudolph & Forsterling, 1997) verbs should show biases. The validity of these predictions is unknown, as to date our knowledge of IC has been based largely on a limited number of verbs selected primarily because they are known to have strong IC pronoun biases. In a thorough meta-analysis, Rudolph and Forsterling (1997) found that although dozens of studies had been conducted and four different languages had been tested, data on individual verbs had been reported for only 256 verbs, partly because many of the studies employed the same small set of verbs (but see discussion of Ferstl et al., 2011, in Appendix 3). This narrow focus increases the risk of over-fitting the data. Thus, in Experiment 1, verbs were chosen solely on the basis of frequency and their ability to have both animate subjects and objects. We use these data to evaluate the verb classes proposed by Brown and Fish (1983a), Rudolph and Forsterling (1997), and the Linguistic Category Model (Semin & Fiedler, 1998, 1991), and then examine whether any of the classes proposed by VerbNet are systematic in their bias.

Method

Participants

Participants were recruited and tested online through coglanglab.org. We analysed the data from the 1,365 participants who were native English speakers, reported not

TABLE 1
Verb classes employed in earlier semantic structure accounts

Taxonomy	*Verb classes*	*Predicted bias*	*Examples*
Brown and Fish	Experiencer-stimulus	Object	detects, admires
	Stimulus-experiencer	Subject	calms, confuses
	Action	Subject	loses, cools
Au	Experiencer-stimulus	Object	detects, admires
	Stimulus-experiencer	Subject	calms, confuses
	Action-patient	Object	criticises, cools
	Action-agent	Subject	loses, hits, calls
Rudolph & Forsterling/McKoon et al.	Experiencer-stimulus	Object	detects, admires
	Stimulus-experiencer	Subject	calms, confuses
	Agent-evocator	Object	criticises, scorns
	Agent-patient	Subject	loses, hits, calls
Linguistic category model	Experiencer-stimulus	Object	detects, admires
	Stimulus-experiencer	Subject	calms, confuses
	Descriptive action verb	None/Weak Subject	Kiss, punch
	Interpretive action verb	Subject	help, break

having participated in the experiment previously, and were over 10 years old ($M = 29.9$, $SD = 13$).

Materials and procedure

As in previous studies, in order to create sentences with ambiguous pronouns, we selected verbs that allow both animate subjects and objects. Because we wanted a sample that was both as representative and unbiased as possible, this last criterion was applied loosely, leading to the inclusion of a number of verbs that are marginal with both animate subjects and objects, or allow them only when the verb is used in a dispreferred sense (see Appendixes 1–2). We chose the 720 most common English verbs (Frances & Kucera, 1982) that met these criteria (Appendix 1).

Each subject was tested on 25 verbs randomly sampled from the total set. In order to minimise the effects of words other than the verb on the judgments, all sentences were of the form *Sally VERBs Mary because she is a dax*. An example trial is presented below:

(12) Sally frightens Mary because she is a dax.
Who do you think is the dax?
Sally Mary

The participant indicated his/her choice by clicking one of the names with the mouse. The order of the names (e.g. Sally, Mary) was randomised on each trial, with the grammatical subject sometimes being listed on the left and sometimes on the right. Participants were told a dax is a type of person but given no more information. The same novel word was used for all trials; randomisation of items across participants should mitigate any systematic order effects. The names (e.g. *Sally* and *Mary*) were chosen randomly without repetition for each participant from a list of common American female names taken from a recent census. All sentences were presented visually.

Participants were given two example sentences and encouraged to recognise the ambiguity of the pronoun (e.g. *Sally helps Mary because she is a dax. You might think that daxes are very helpful and that Sally is the dax. Otherwise, you might imagine that in this story daxes deserve help and that Mary, a dax, gets help from Sally*).

Taxonomies

Many verbs have multiple meanings. For instance, *Mary touches Sally* can be interpreted as a contact event or as a psychological event. This not only complicates classifying the verb for taxonomic analysis, but introduces noise as different participants may interpret the verb differently, thus arriving at different IC biases. (Interestingly, despite the ubiquity of verbal polysemy, this issue does not appear to have been addressed previously in the IC literature.) Thus, for all analyses involving verb taxonomies, we excluded all polysemous verbs, operationalised as verbs that fall into more than one VerbNet class, unless the sentence frame that was used in the study ruled out consideration of the additional class (e.g. because it requires an intransitive frame or an inanimate object), leaving us with 328 monosemic verbs, which we classified according to four well-known IC verb taxonomies:

Brown & Fish (1983a). The Brown and Fish taxonomy consisted of two types of psych verbs (experiencer-stimulus verbs and stimulus-experiencer verbs) and action

verbs. Brown and Fish (1983a) do not give explicit definitions of these classes. Based on examples given by Brown and Fish, we assume psych verbs include all verbs of emotion (*fear, frighten*), cognition (*know, understand*), and perception (*see, hear*). Transitive psych verbs are easily identified in VerbNet: classes 30.1 (*detect, hear*), 30.2 (*discover, recognise*), 31.2 (*fear, love*), 32.1 (*covet, crave*), 87.2 (*misinterpret, misunderstand*) are experiencer-stimulus and class 31.1 (*frighten, surprise*) is stimulus-experiencer. All other verbs were classified as action verbs.

Rudolph & Forsterling (1997)/McKoon et al. (1993). Rudolph and Forsterling (1997) adopt the Brown and Fish classifications of experiencer-stimulus and stimulus-experiencer verbs, but divide action verbs into agent-evocator and agent-patient verbs based on their IC bias (the former are defined as object biased; the latter, subject biased). To make this a noncircular taxonomy, we follow Ferstl et al. (2011) in classifying the examples of agent-evocator verbs given by Rudolph and Forsterling (1997) according to VerbNet and then tagging all verbs in those classes as agent-evocator. The intuition here is that on any *semantic structure* account, verbs with similar meanings should have similar biases (and verbs within one VerbNet class have highly similar meanings). The only verb class exemplified by at least one monosemic verb was class 33 (*praise, slander*) (for polysemous verbs, it is impossible to know which meaning the authors intended to invoke). All other action verbs were classified as agent-patient. This results in a taxonomy identical to that proposed by McKoon et al. (1993).

It should be noted that Rudolph and Forsterling (1997) also include class 33 verbs in their list of agent-patient verbs (*slander*), and thus by the logic used above this class is *both* agent-patient and agent-evocator. However, the model with class 33 verbs as agent-evocator performs the best, and so that is the only one discussed below. Note that McKoon et al. (1993) specifically single out class 33 verbs (*praise, slander*) as object biased, and thus for them the problem does not arise.

Au (1986). Au's taxonomy is for most purposes identical to Rudolph and Forsterling's. There are two classes of action verbs: action-agent, defined as subject-biased action verbs, and action-patient, defined as object-biased action verbs (again, psych verbs are handled identically to Brown and Fish). The primary difference is that Au provides more examples of verbs in the two action verb classes: all monosemic action-patient examples were either in VerbNet class 33 (*praise, slander*) or class 45.4 (*cool, improve*). As was the case with Rudolph and Forsterling, verbs from both these classes are also included in the subject-biased action-agent class. However, attempting to account for this leads to taxonomies that perform even more poorly and thus are not discussed further here.

Linguistic Category Model. The most comprehensive description of the Linguistic Category Model appears in Semin and Fielder (1991), who identify four types of verbs: *state verbs, state action verbs, descriptive action verbs*, and *interpretive action verbs*. The first two classes appear to be identical to experiencer-stimulus and stimulus-experiencer verbs, respectively [the former "refer to mental and emotional states; no clear definition of beginning and end; do not readily take progressive forms; not freely used in imperatives" (p. 5), and the latter describe an "implicit action frame by the sentence subject that leads to the experience of a state in the object of the sentence" (p. 6)]. The remaining verbs are either descriptive action verbs, which entail that the action have one physically invariant feature, or interpretive action verbs, which do not. The authors seem to have in mind a fairly liberal notion of what is "physically

invariant". Examples of such verbs include *meet, summon, stop*, *prepare*, *visit*, and *wake up* (Semin & Fiedler, 1988, 1991; Semin & Marsman, 1994). Interpretive action verbs are predicted to be subject biased, whereas descriptive action verbs are predicted to be nonbiased. In order to avoid experimenter bias, we asked naive participants to code verbs according to the definitions described above.[5]

Results

We first provide a descriptive overview of the results. We then evaluate the predictions of each of the previous taxonomies. Finally, we evaluate a new taxonomy based on VerbNet.

Each verb was evaluated by an average of 47 participants. The distribution of results is shown in Figure 1, and results for each verb are presented in Appendix 1. Across all 720 verbs, there was a slight overall bias in favour of choosing the object as the referent of *she*, overall object bias: 59.2%, $SE = 0.6\%$, $t(719) = 15.93$, $p < .0001$. A total of 37 verbs exhibited a significant subject bias (p's $\leq .05$) and 265 a significant object bias (p's $\leq .05$). Of these, 3 subject-biased and 93 object-biased verbs survive a conservative Bonferroni correction for 720 comparisons. Thus, the bulk of the verbs tested showed no strong bias, a fact confirmed by Hartigan dip test analyses (Hartigan & Hartigan, 1985) implemented in R (Maechler, 2009; R Development Core Team, 2009), which found no evidence of a bimodal distribution in IC bias (.009, $p > .9$). This is consistent with previous reports that many verbs do not elicit a systematic pronoun bias (Garvey et al., 1974; Semin & Fiedler, 1991).

We investigated whether this unimodality was due to noise in the distribution caused by polysemous verbs. After excluding all verbs with more than one possible use, the remaining 328 monosemic verbs again showed a broad distribution of biases (Figure 2) with a slight overall object bias, 58.4% choosing object, $SE = 0.9\%$, $t(327) = 9.48$, $p < .0001$. Again, there is no evidence of a bimodal distribution, Hartigan's $dip = 0.015$, $p > .6$, again indicating that a significant proportion of the verbs showed no clear IC pronoun bias. This pattern is not attributable to the inclusion of verbs that are marginal for two animate arguments; many verbs that typically take two animate arguments (*troubles, commands, teaches*) showed no clear IC bias (see also Appendix 1).

These results are problematic for previous taxonomies on which most or all verbs should exhibit IC biases (Table 1). We considered the IC biases for each of the classes

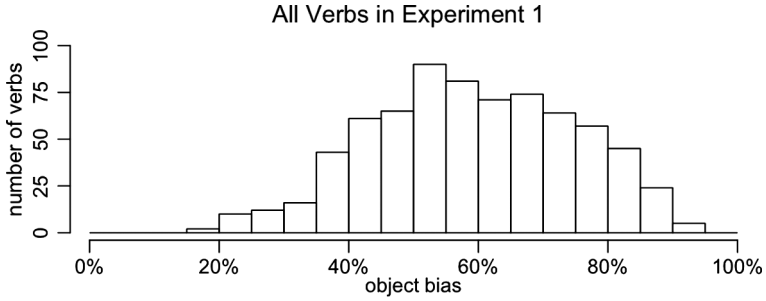

Figure 1. Histogram of object biases for the 720 verbs in Experiment 1.

[5]The 328 monosemic verbs (see below) were divided into eight lists. A total of 12 English-speaking participants recruited through Amazon Mechanical Turk participated in each list, with 23 excluded for failing to follow directions.

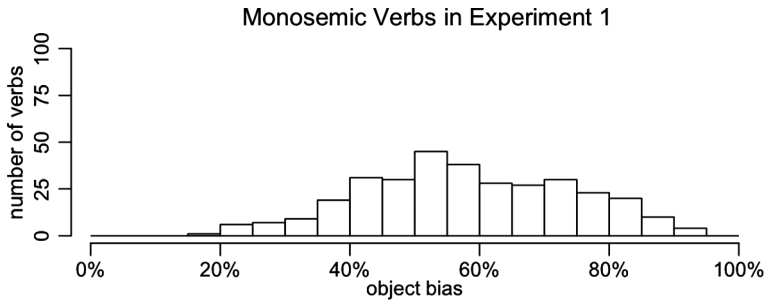

Figure 2. Histogram of object biases for the 328 monosemic verbs in Experiment 1.

in the four previous taxonomies. Because we are primarily interested in knowing whether verb class predicts IC bias better than the grand mean, and the grand mean in this experiment was object biased, we compared the IC biases in each class to the grand mean for monosemic verbs (58.4%). While analyses based on these taxonomies confirmed that stimulus-experiencer psych verbs are subject-biased and experiencer-stimulus psych verbs are object biased—a prediction common to Brown and Fish (1983a), Rudolph and Forsterling (1997), and the Linguistic Category Model (Semin & Fiedler, 1988, 1991)—these taxonomies do not isolate nonpsych verbs which are subject biased, and all but the Rudolph and Forsterling/McKoon et al. taxonomy fail to isolate nonpsych verbs which are object biased (Table 2). This is despite the fact that many such verbs do have significant biases (Appendix 1).

In the case of the Linguistic Category Model, we considered the possibility that naive participants did not have the metalinguistic knowledge to accurately distinguish descriptive action and interpretive action verbs. The first author recoded all 328 monosemic verbs twice: first with a relatively strict interpretation of Semin and Fiedler's (1991) definition of descriptive-action verbs, and second with a broader interpretation so as to include verbs of communication (cf. *summon* and *call*; Semin &

TABLE 2
Results for the four previous semantic structure accounts by verb class, with class object bias mean (standard deviation), compared to grand average for monosemic verbs. Note that all four employ the same experiencer-stimulus and stimulus-experiencer classes

		Object bias	
Class	N	*Diff. from Mean*	*Significance*
Brown & Fish, Rudolph & Forsterling/McKoon et al., Au, and Linguistic Category Model			
Experiencer-stimulus	32	+10% (17%)	$t = 3.40$, $p = .002$
Stimulus-experiencer	16	−19% (13%)	$t = 5.77$, $p = .00004$
Brown and Fish			
Agent-patient	280	0% (15%)	$t < 1$
Rudolph & Forstering/McKoon et al.			
Agent-evocator	18	+12% (13%)	$t = 3.95$, $p = .001$
Agent-patient	262	−1% (15%)	$t < 1$
Au			
Agent-evocator	46	0% (16%)	$t < 1$
Agent-patient	234	0% (15%)	$t < 1$
Linguistic Category Model			
Descriptive action	19	4% (14%)	$t = 1.42$, $p = .17$
Interpretive action	261	0% (15%)	$t < 1$

Fiedler, 1988). On the former (strict) taxonomy, again neither class was significantly biased. On the latter (broad) taxonomy, interpretive action verbs were again nonbiased ($N = 214$, $M_{\text{diff}} = -1\%$, $SD = 16\%$, $t = 1.07$, $p = .29$), while descriptive action verbs were significantly object biased ($N = 66$, $M_{\text{diff}} = +4\%$, $SD = 13\%$, $t = 2.31$, $p = .02$), in both cases contrary to predictions (Table 1). Note that even this latter effect was small, and only a slight majority of the verbs were numerically more object biased than the grand mean (38 of 66).

Thus, when tested against a representative sample of verbs, none of the previous taxonomies discussed successfully picked out both subject- and object-biased action verbs.

VerbNet verb class analyses

We classified the 328 monosemic verbs by VerbNet verb class in order to determine whether this more fine-grained analysis would reveal coherent subclasses with consistent IC biases. IC bias was modelled with a bi-directional stepwise linear regression, with any verb class for which at least five verbs were tested included as predictors. There were 11 such classes covering 135 of the 328 verbs. The resulting model ($R^2 = 0.2$) identified six verb classes, all of which exhibited significant biases (Table 3). The biases for the remaining five classes with at least five verbs were small ($< 5\%$) and did not approach significance ($ts < 1$). These classes were class 13.2 (*lose, relinquish*), class 13.5.1 (*attain, buy*), 13.5.2 (*accept, obtains*), class 30.1 (*detect, hear*), and class 48.1.2 (*define, exhibit*).

Comparison of taxonomies

We next compare the accuracy of the predictions for the four taxonomies. Unlike the previous taxonomies, the VerbNet taxonomy makes no predictions about directionality (but see General Discussion); however it does predict that all verbs in a class should pattern similarly. Thus, a verb was counted as conforming to predictions if its bias was the same as for the class as a whole (subject-, object-, or neither). Verbs were counted as biased if $p < .10$ in a two-sided binomial test.[6] The VerbNet taxonomy correctly classified 56% of verbs (excluding the 193 verbs for which the taxonomy currently makes no predictions). The four previous taxonomies only correctly

TABLE 3
VerbNet verb classes for which at least five verbs were tested, with class object bias mean (standard deviation), compared to grand average for monosemic verbs

		Object bias		
Class	N	*Diff. from Mean*	*Significance*	*Examples*
30.2	6	+10% (8%)	$t = 2.77$, $p = .04$	discovers, recognises, watches
31.1	16	−19% (13%)	$t = 5.77$, $p = .00004$	calms, confuses, frustrates, troubles
31.2	17	+17% (15%)	$t = 4.54$, $p = .0003$	admires, cherishes, despises, loves
33	18	+12% (13%)	$t = 3.95$, $p = .001$	blames, congratulates, thanks
45.4	28	−7% (12%)	$t = 3.32$, $p = .003$	cools, dries, improves, revives
59	6	−12% (7%)	$t = 4.35$, $p = .007$	compels, dares, fools

[6] Note that the Linguistic Category Model was credited with predicting no bias for descriptive action verbs. Its performance drops if it is considered to have predicted subject biases for those verbs.

classified from 28% to 31% (Table 4), primarily due to predicting subject biases for many nonbiased or even object-biased verbs (Table 5).

Chance performance for each of the taxonomies was estimated using Monte Carlo simulation. The results for the verbs were randomly permuted 10,000 times, holding each taxonomy's predictions constant. Thus chance performance could be estimated while accounting for the fact that each taxonomy made different predictions about the base rates of subject and object biases.

All taxonomies performed significantly above chance ($ps < .05$). For most of the taxonomies, this success was attributable entirely to correctly predicting the biases for classes 31.1 (*frighten, surprise*) and 31.2 (*fear, love*), about which all taxonomies agree. With those words excluded, only the Rudolph and Forsterling/McKoon et al. and the VerbNet taxonomies performed above chance ($ps < .001$), with the latter still performing considerably better overall (Table 4).

Discussion

The four previous taxonomies fared quite poorly when tested against a representative sample of verbs. In particular, as they predict that most or all verbs should have significant IC biases, whereas most do not. These taxonomies cannot be much improved by changing the predictions from some of the verb classes from "biased" to "nonbiased", as the taxonomies would still fail to pick out biased action verbs, of which there are many. These taxonomies do better on psych verbs but remain far from perfect.

In contrast, the VerbNet analyses pick out four classes of nonbiased action verbs [classes 13.2 (*love, relinquish*), 13.5.1 (*attain, buy*), 13.5.2 (*accept, obtain*), and 48.1.2 (*define, exhibit*)] and three classes of biased action verbs [classes 33 (*praise, slander*), 45.4 (*cool, improve*), and 59 (*compel, dare*)], more accurately capturing the pattern of results. Moreover, the results for the biased classes are remarkably uniform: in class 33 (*praise, slander*), 14 of 18 verbs were numerically object biased and in class 59 (*compel, dare*) all six verbs were numerically subject biased. Class 45.4 (*cool, improve*) was less consistent, with 18 of 28 verbs numerically subject biased. It should be noted that many verbs in this class are only marginally acceptable with two animate arguments (e.g. *improves*), which may have contributed to its unreliability. It should also be noted that while VerbNet largely agrees with the previous taxonomies in terms of psych verbs, with experiencer-stimulus verbs being object-biased and stimulus-experiencer verbs being subject biased, there was one class of experiencer-stimulus verbs for which there was no strong evidence of bias, class 30.1 (*detect, hear*). Given that only a small

TABLE 4

Chance and observed percentage of verbs conforming to predictions for each of the five taxonomies, both across all monosemic verbs (328 total) and excluding classes 31.1 and 31.2 (295 total). The cutoff for a significant bias was $p = .10$. Percentages for VerbNet were calculated only out of verbs for which it makes predictions (135 monosemic, 102 excluding classes 31.1 and 31.2)

	All monosemic verbs		*Excluding class 31.1, 31.2*	
	Chance (%)	*Observed (%)*	*Chance (%)*	*Observed (%)*
Brown and Fish	25	28	22	22
Rudolph & Forsterling/McKoon et al.	25	31	23	26
Au	25	28	23	22
Linguistic Category Model	26	31	22	22
VerbNet	33	56	36	49

TABLE 5
Percentage (number of) verbs that conformed to predictions for each of the four taxonomies. The cutoff for a significant bias was $p = .10$

	Subject bias	*Object bias*	*No bias*
Predicted subject bias			
Brown and Fish	25% (74/296)	21% (63/296)	54% (159/296)
R & F/McKoon et al.	26% (73/278)	52% (52/278)	55% (153/278)
Au	12% (62/250)	21% (52/250)	54% (136/250)
Linguistic Category Model	26% (72/277)	21% (57/277)	53% (148/277)
VerbNet	54% (27/50)	0% (0/50)	46% (23/50)
Predicted object bias			
Brown and Fish	22% (7/32)	56% (18/32)	22% (7/32)
R & F/McKoon et al.	16% (8/50)	58% (29/50)	26% (13/50)
Au	24% (19/78)	37% (29/78)	38% (30/78)
Linguistic Category Model	22% (7/32)	56% (18/32)	22% (7/32)
VerbNet	7% (3/41)	66% (27/41)	27% (11/41)
Predicted no bias			
Brown and Fish	NA	NA	NA
R & F/McKoon et al.	NA	NA	NA
Au	NA	NA	NA
Linguistic category model	(2/19)	(6/19)	58% (11/19)
VerbNet	25% (11/44)	27% (12/44)	48% (21/44)
No predictions			
Brown and Fish	NA	NA	NA
R & F/McKoon et al.	NA	NA	NA
Au	NA	NA	NA
Linguistic Category Model	NA	NA	NA
VerbNet	21% (40/193)	22% (42/193)	58% (111/193)

number of verbs in that class were tested (six), one cannot rule out the possibility of sampling/measurement error, but the possibility of finer-grained distinctions between experiencer-stimulus psych verbs merits further research.

EXPERIMENT 2

The above results demonstrate that IC bias varies systematically with respect to the fine-grained verb classes identified by VerbNet. But the semantic structure hypothesis makes a stronger prediction. To the extent that IC biases are caused solely by differences in verb semantics, we should expect all verbs in a given class to show similar IC biases. The existing data sets are not well suited for testing this prediction. Most studies have used a small set of verbs that were specifically selected because they were believed to have a strong object or subject bias (Rudolph & Forsterling, 1997). Experiment 1 avoids the problem of selection bias, but is not well designed to test the consistency of IC biases within a given verb class for two reasons. First, the study is exploratory in the sense that the bias of each class was determined empirically based on the behaviour of its members (rather than predicted *a priori*). Second, most verb classes contained fewer than 20 verbs and thus minimal information was available about the distribution of verbs within a class. Finally, on average fewer than 50 judgments were collected for each verb. As the judgments were binary, this limited resolution plus sampling error means that our estimates of individual verbs' biases were relatively imprecise, potentially smearing the distribution of IC biases within any given class. Thus, in Experiment 2, we collect substantially more judgments per verb in

order to minimise measurement error and sample a much larger number of verbs within a given class to test consistency.

Specifically, in Experiment 2, we collected IC judgments on all the verbs in class 31.1 (*frighten, confuse*) and 31.2 (*fear, love*) that were listed in Levin (1993). There were three reasons for selecting these two classes for further analysis. First, both are large classes thus providing us with sufficient verbs to provide a strong test of within-class uniformity. Second, unlike the verbs in some of the other classes [e.g. 45.4 (*cool, improve*)], these verbs are readily used with both an animate subject and an animate object, resulting in more natural stimuli. Third, these classes make up the bulk of transitive psych verbs and psych verbs have been played a central role in both IC research and in the study of argument realisation (Au, 1986; Brown & Fish, 1983a, 1983b; Dowty, 1991; Ferstl et al., 2011; Goikoetxea, Pascual, & Acha, 2008; Greene & McKoon, 1995; Jackendoff, 1990; Levin, 1993; Levin & Rappaport Hovav, 2005; Pesetsky, 1995; Pinker, 1989; Rudolph & Forsterling, 1997; Semin & Fiedler, 1988, 1991; Semin & Marsman, 1994; Talmy, 2000a, 2000b). While it is widely believed that class 31.1 (*frighten, surprise*) verbs are uniformly subject biased and class 31.2 verbs (*fear, love*) are uniformly object biased, only a handful of such verbs have been tested and whether the results are truly general is unknown.

Psych verbs are problematic for theories of argument realisation because these verbs appear to involve the same thematic roles but they vary in how they express them (experiencer-stimulus or stimulus-experiencer). Within IC research there is a robust consensus that the IC bias of psych verbs follows the STIMULUS (Brown & Fish, 1983a, 1983b; Rudolph & Forsterling, 1997); nonetheless non-negligible numbers of exceptions have been reported (Ferstl et al., 2011; Goikoetxea et al., 2008). And indeed, in Experiment 1 we found one class of experiencer-stimulus verbs [30.1 (*detect, hear*)] which showed no systematic IC bias. Thus we cannot assume that biases are consistent within the class 31.1 (*frighten, surprise*) or class 31.2 (*fear, love*). In addition, some theorists have proposed that the variability in argument realisation in psych verbs reflects differences in their meaning which could potentially influence their IC bias. For example, Pesetsky has suggested that the subject of a *frighten* verb is actually the CAUSE of the emotion while the object of a *fear* verb is merely the TARGET of this emotion. This might lead us to expect strong consistency across verbs among class 31.1 (*frighten, surprise*) but more variability within class 31.2 (*fear, love*) verbs.

Method

Participants

Participants were recruited and tested online through coglanglab.org. We included only the 1,025 participants (mean age: 30, $SD = 13$) who completed the experiment were native English speakers, and reported not having participated in the experiment previously.

Materials and procedure

There are considerably more class 31.1 (*frighten, surprise*) verbs (220) than class 31.2 (*fear, love*) verbs (44). Thus, if stimuli were randomised, participants would see primarily class 31.1 verbs, which by hypothesis are subject biased. This risks priming subject resolution, distorting the phenomenon of interest. Thus, each participant was presented with 12 verbs randomly selected without replacement from each of two verb

classes, resulting in an average of 56 judgments for each *frighten* verb and 280 judgments for each *fear* verb.

Results and discussion

Analyses below use the grand mean of 58.4% object bias as our conservative "chance" threshold. Individual verb results are presented in Appendix 2; the distributions are shown in Figure 3. In contrast to Experiment 1, the distribution was clearly bimodal (Hartigan's $dip = 0.036$, $p < .05$). Class 31.1 (*frighten, surprise*) showed a strong subject bias, 35.7% object bias, $SE = 1.0\%$, $t(219) = 14.86$, $p < .01$, while class 31.2 (*fear, love*) showed a strong object bias, 81.5% object bias, $SE = 1.1\%$, $t(43) = 28.53$, $p < .01$, and the two classes were significantly different from one another $t(262) = 31.27$, $p < .01$.

All class 31.2 (*fear, love*) verbs exhibited object biases, 41/44 significantly so (39/44 after Bonferroni correction), while 202/220 class 31.1 verbs (*frighten, surprise*) exhibited subject biases, 170 significantly so (110 survive Bonferroni correction).[7] Only six of the latter showed significant object biases (1 after Bonferroni correction). Thus, semantic class was a very strong predictor of the IC pronoun bias, consistent with the *semantic structure* account.

As in Experiment 1, we conducted further analyses focusing on monosemic verbs as defined by VerbNet. The resulting distribution was again bimodal (Hartigan's $dip = 0.040$, $p < .02$; Figure 4). Of the remaining 171 class 31.1 (*frighten, surprise*)

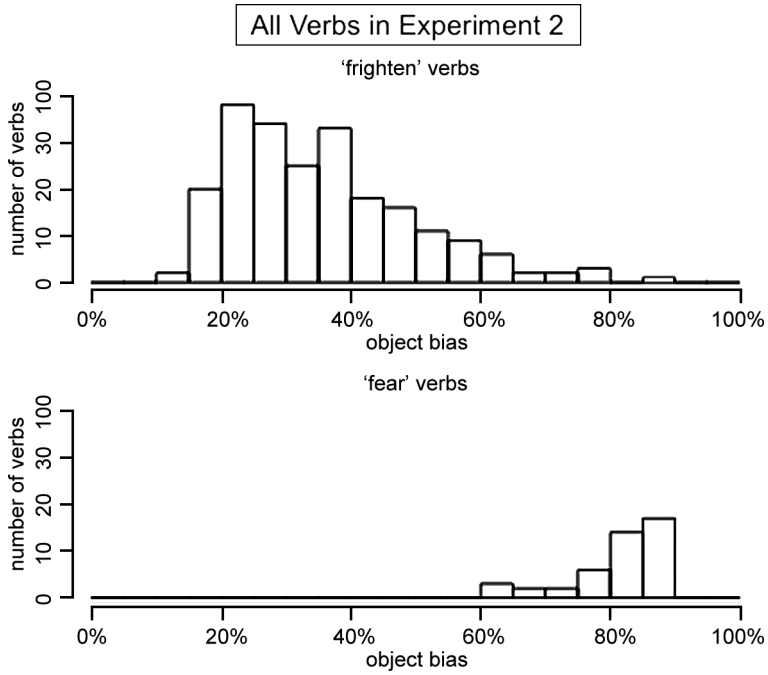

Figure 3. Histograms of biases for all class 31.1 (*frighten, surprise*) and class 31.2 (*fear, love*) verbs in Experiment 2.

[7] All 44 class 31.2 (*fear, love*) verbs were significantly different from the 50% chance threshold, even after Bonferroni correction. Using the 50% threshold necessarily raises the bar for subject-biases. Nonetheless, 184 class 31.1 (*frighten, surprise*) verbs still show a numeric subject bias (130 significantly), 34 an object bias (12 significantly), and 2 no bias.

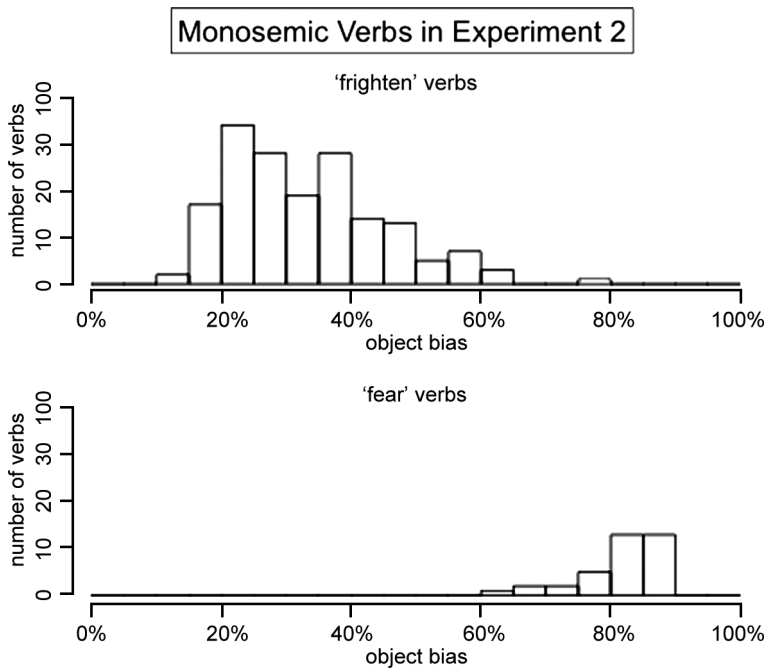

Figure 4. Histograms of biases for monosemic class 31.1 (*frighten, surprise*) and class 31.2 (*fear, love*) verbs in Experiment 2.

verbs, all but 7 (*wounds, dejects, cows, alienates*, *discourages, placates, torments*) exhibited a subject bias, 142 significantly (95 after Bonferroni correction; See Figure 4). Only one class 31.1 *frighten* verb (*alienate*) was significantly object biased and not after Bonferroni correction. All but 1 (*stands*) of the 36 remaining class 31.2 (*fear, love*) verbs were significantly object biased (33 after Bonferroni correction). Moreover, there was no overlap between the distributions of the two classes, with the exception of *alienate*.

Thus, while Experiment 1 and previous studies (e.g. Ferstl et al., 2011; Goikoetxea et al., 2008) found considerable overlap in the distributions for the broadly defined experiencer-stimulus and stimulus-experiencer classes, class 31.1 (*frighten, surprise*) and class 31.2 (*fear, love*) verbs are categorically different in their behaviour. Importantly, we find that this categoricity extends beyond the relatively small number of verbs that have been repeatedly tested.

GENERAL DISCUSSION

Experiments 1 and 2 demonstrate that IC biases in pronoun interpretation vary systematically across semantic classes of verbs that are independently motivated based on patterns of argument realisation. In Experiment 1 and the reanalysis of Ferstl et al. (2011), we investigated 11 different verb classes, finding significant biases 6 of them. In Experiment 2, we investigated two of these classes—class 31.1 (*frighten, surprise*) and class 31.2 (*fear, love*)—finding that the IC bias was consistent for the vast majority of members in both classes. We also found converging results from a reanalysis of data from 305 verbs reported by Ferstl et al. (2011) (see Appendix 3). These findings suggest that IC bias varies systematically with coherent, independently defined verb

classes. These results also demonstrate why previous IC verb taxonomies have shown inconsistent results: these taxonomies collapsed across different sets of verbs that exhibit systematically different biases. Indeed, we find when applied to large sets of verbs, these older taxonomies are close to or at chance in predicting IC bias.

These results are fully consistent with the *semantic structure* hypothesis, which directly predicts a systematic relationship between semantic structure and IC bias. Our findings are not directly predicted or explained by alternate accounts, such as the *arbitrary semantic tag* or *world knowledge* accounts. However, these hypotheses could be amended or extended to account for these findings, thus below, we explore how these alternative accounts are constrained by the present data. First, however, we discuss theories of semantic structures of verbs that could potentially support causal inferences and consider whether IC bias is continuous or categorical.

THE SEMANTICS OF VERBS

VerbNet verb classes are defined syntactically but are argued to represent coherent semantic classes (cf. Levin & Rappaport Hovav, 2005). We suggested above that the semantics of these classes signal—directly or indirectly— who caused the event or state described by the verb. This information, in combination with expectations about the content of the subordinate clause introduced by *because* (Brown & Fish, 1983a; Garvey & Caramazza, 1974; Garvey et al., 1974; Kehler et al., 2008), would provide a straightforward representational basis for the observed pronoun bias.[8] In this section, we describe two frameworks that have been proposed for representing the semantics of verbs (thematic roles and predicate decomposition, described below) and discuss how each might explain the relevant data. This is an active area of research: there are many competing thematic role and predicate decomposition theories and consensus on the correct description of verbal semantics is a long way off. Thus our goal in this section is simply to describe whether and how such theories could, in principle, account for IC bias. Nonetheless, it is easiest to describe classes of theories by outlining specific examples. The most fully implemented thematic role theory and most fully implemented predicate decomposition theory are both found in VerbNet.

Thematic roles

Thematic roles are invoked in linguistic theory to help explain how the different arguments of the verb are syntactically encoded in a clause (for review, see Levin & Rappaport Hovav, 2005). For example, thematic roles are invoked to explain which argument of a two-place predicate will surface as the syntactic subject (e.g. in *Sally broke the vase*, Sally is an AGENT and AGENTS are mapped onto subject position). Starting with Brown and Fish (1983a), several previous *semantic structure* accounts invoked thematic roles to explain IC: namely, some thematic roles are inherently causal (e.g. AGENT, STIMULUS), and thus comprehenders expect entities filling those roles (like Sally in *Sally frightened Mary*) to be the "implicit" cause of the event in question. The data and analyses above suggest that these theories were insufficiently

[8] Early discussions of IC implicitly assumed that a subordinate clause introduced by *because* necessarily encodes the cause of the event in the main clause. Recently, several authors have suggested that *because* introduces an explanation, rather than a cause per se (Kehler, 2002; Kehler et al., 2008; Pickering & Majid, 2007). Either account is consistent with the analysis here, since explanations by necessity are more likely to refer to entities that were causally responsible for an event (Kehler, 2002; Kehler et al., 2008).

nuanced to capture patterns in IC biases. Interestingly, modern thematic role theorists have typically found it necessary to posit far more than the 4–5 thematic roles employed in previous IC theories in order to account for differences in verb semantics and syntactic behaviour. VerbNet, for example, utilises 33 different thematic roles. Perhaps this more extensive set of thematic roles will be able to capture IC bias patterns.

We explored this possibility in two ways. First, we analysed the monosemic verbs from Experiment 1, coded for VerbNet thematic roles, excluding 68 verbs for which VerbNet suggested more than one set of thematic roles.[9] Each thematic role was entered into a bi-directional stepwise linear regression composed of main effects only with each thematic role as a predictor, coded as "1" if the thematic role appeared in subject position, as " − 1" if it appeared in object position, and "0" otherwise. The resulting model contained three thematic roles: EXPERIENCER (equivalent to Brown and Fish's (1983a) thematic role of the same name), PRODUCT (an entity that is created during the event, as in the object of *design*, *rationalise* or *rebuild*), and STIMULUS (the object of cognition and perception verbs—note that this is a small subset of the relevant verbs in Brown and Fish's (1983a) taxonomy), all of which were significant predictors of bias ($ps < .05$).[10] Surprisingly, all predicted that the pronoun would be resolved to the *other* argument. Thus, if there are thematic roles which always attract pronoun resolution, this method combined with this particular thematic role theory could not identify them.

However, it may be that IC bias cannot be predicted directly from the thematic role borne by an argument but by the causal strength of that thematic role relative to the thematic role borne by the other verbal argument. We investigated this possibility in our second analysis. We used the Batchelder − Bershad–Simpson scaling method (Batchelder, Bershad, & Simpson, 1992) to estimate a hierarchy for the thematic roles investigated above (Figure 5).[11] This statistical technique has been widely used to estimate dominance hierarchies in social animals based on the outcomes of dyadic interactions (Jameson, Appleby, & Freeman, 1999) and is based on a method introduced for ranking chess players (Elo, 1978). Note that an advantage of this method is it does not require two thematic roles to have actually appeared with the same verb in order to estimate which is more highly ranked. Thus, CAUSE is ranked higher than STIMULUS not because CAUSE-STIMULUS verbs are known to be biased towards the patient (no such verbs exist), but because CAUSE-EXPERIENCER verbs are strongly biased towards the cause [these are the class 31.1 (*frighten, surprise*) verbs], whereas EXPERIENCER-STIMULUS verbs are only weakly biased towards the STIMULUS [these are class 30.1 (*detect, hear*) and 30.2 (*discover, recognise*) verbs]. Note that, unlike in the IC literature, VerbNet does not classify the nonEXPERIENCER argument of class 31.2 (*fear, love*) verbs as a STIMULUS; rather, this is a THEME, a categorisation typical in linguistics (see Levin & Rappaport Hovav, 2005). This explains the relatively high ranking of THEME, as such verbs are strongly biased in favour of the THEME.

[9] An example is the verb *dry* (class 45.4). In *Bill dried the clothes*, VerbNet codes *Bill* as an AGENT and *the clothes* as a PATIENT. In *The hairdryer dried the clothes*, VerbNet codes *the hairdryer* as an INSTRUMENT and *the clothes* as a PATIENT. Note that the issue here is not purely animacy: animate beings can be used as instruments (*John wiped the floor with Bill*).

[10] One additional thematic role (THEME) was retained in the stepwise regression but was not itself a significant predictor, $t(160) = 1.53$, $p = .13$.

[11] Additionally, we excluded 12 verbs, discussed further below, for which both arguments bore the same thematic role. A total of 235 verbs remained.

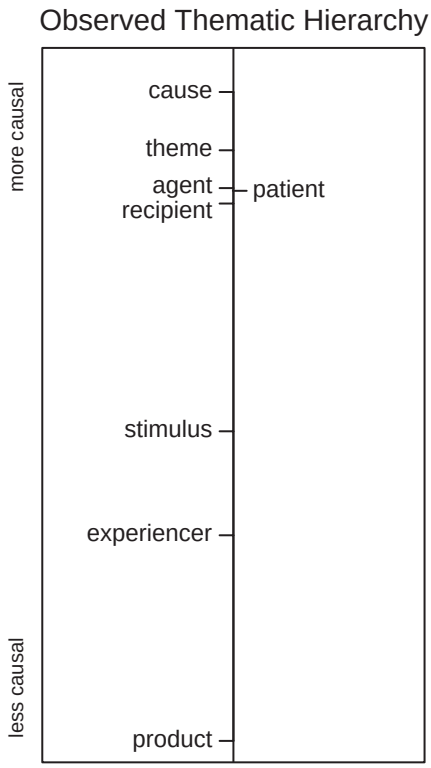

Figure 5. Thematic roles on a causal hierarchy, as estimated by the Batchelder-Bershad-Simpson scaling method.

The resulting thematic hierarchy (Figure 5) successfully predicted the numeric direction of bias for 60% of the relevant verbs—far better than the more coarse-grained thematic role theories employed by Brown and Fish (1983a), Au (1986) or Rudolph and Forsterling (1997), and on par with the VerbNet verb classes, despite having considerably fewer degrees of freedom than the latter. This fit would likely improve if we allowed that verbs for which the arguments are in similar positions in the hierarchy should be nonbiased. We leave implementation of this for future research. We further note that much of the hierarchy is intuitively correct, with CAUSE being the most causal and PRODUCT—an entity which does not exist prior to the event and thus cannot have caused the event—being the least causal.

Nonetheless, there are some limitations to this hierarchy. One would expect that if both arguments of the verb bore the same thematic role, the verbs would be unbiased. However, the seven verbs (*better, dominate, exceed, generate, overcome, own,* and *possess*) which had a THEME in both subject and object position were on average subject biased, $t(6) = 3.44$, $p = .01$. Similarly, the five verbs (*divorce, embrace, fight, marry,* and *visit*) listed with ACTOR in both positions were on average object biased, $t(4) = 2.99$, $p = .04$. Similarly, not all of the rankings on the hierarchy can be clearly independently motivated. That is, there is no clear reason that RECIPIENT should be less causal than EXPERIENCER. It should be reiterated that the hierarchy in Figure 5 is a function not just of our data but of the thematic role theory implemented by VerbNet. Whether other thematic role theories can successfully address these limitations is a question for further research.

Predicate decomposition

Predicate decomposition theories have emerged in part to address limitations in the explanatory power of thematic role theories, such as the fact that neither the exact list of thematic roles nor the fact that many possible thematic role combinations in any given language are unattested (English has no CAUSE-STIMULUS verbs) is left unexplained and must be stipulated (Levin & Rappaport Hovav, 2005). Such theories decompose the semantics of a verb into more primitive predicates, and it is these predicates that assign functions to the arguments of the verb (e.g. Harley, 1995; Jackendoff, 1990; Moens & Steedman, 1988; Pinker, 1989; see Levin & Rappaport Hovav, 2005, for a thorough discussion). For instance, when used transitively, verbs from class 31.1 (*frighten, surprise*) might be decomposed as:

(13) class 31.1: NP1 V NP2
cause(NP1, E) emotional_state(result(E), emotion, NP2)

That is, such verbs have two arguments (NP1 and NP2) and describe two things: (1) NP1 causes event E, and (2) an emotional state is experienced by NP2 as a result of event E. In *Sally frightened Mary*, Sally is NP1, Mary is NP2, and fright is the EMOTION. Depending on the account, NP1 is realised as the sentential subject because it is the argument of CAUSE, because it is the least-embedded argument in the semantic structure, or for some other reason (Levin & Rappaport Hovav, 2005), but on any such account *Sally frightened Mary* literally means that Sally caused Mary to be afraid, straightforwardly predicting a subject IC bias. Importantly, on this theory, all verbs in a class share the same semantic structure. The only difference between members of class 31.1 (*frighten, surprise*) is what type of EMOTION is entered into this semantic structure. Thus, all verbs in a class should have the same IC bias.

VerbNet suggests the following structures for class 45.4 (*cool, improve*) and class 59 (*compel, dare*), the other two subject-biased VerbNet classes identified above:

(14) class 45.4: NP1 V NP2
cause(NP1, E) state(result(E), endstate, NP2)

(15) class 59: NP1 V NP2
force(during(E), NP1, NP2,?Proposition)

In the case of class 45.4 (*cool, improve*), once again the subject of the verb is marked as the cause of the event described by the verb, thus accounting for the IC bias.[12] In the

[12] An individual verb may be able to appear in multiple semantic structures. For instance, class 45.4 is also compatible with the following semantic structure, which is linked to an intransitive syntactic frame:

(16) NP1 V
STATE(RESULT(E), ENDSTATE, NP1)

Thus, the syntactic frames that a verb can appear in reduces to the semantic structures a verb is compatible with, which explains—on this account—why identifying verb classes with syntactic criteria leads to semantically coherent verb classes. A challenge for such theories, then, is to fully account for why each verb is compatible with certain semantic structures and not others (e.g. Rappaport Hovav & Levin, 1988; Pinker, 1989).

case of class 59 (*compel, dare*), the subject is marked as forcing the object to adopt a proposition or state, and is thus similarly causal.

The object-biased class 31.2 (*fear, admire*) has a very different structure. While some theorists have suggested that the object of class 31.2 (*fear, love*) verbs is also a cause (Jackendoff, 1990; Pinker, 1989), VerbNet does not analyse it in this way (see also Pesetsky, 1995). Instead the structure it provides treats the emotional state as arising "in reaction to" the verb's object (17). If we assume that *react to* is the inverse of *cause to*, then we can conclude that the object of *fear* has semantic properties that are similar to the subject of *frighten*, explaining the IC bias:

(17) class 31.2: NP1 V NP2
emotional_state(E, emotion, NP1) in_reaction_to(E, NP2)

The object-biased class 30.2 (*discover, recognise*) similarly involves the IN_REACTION_TO component:

(18) class 30.2: NP1 V NP2
perceive(during(E), NP1, NP2) in_reaction_to(E, NP2)

Levin (1993) suggests a similar structure for class 33 (*praise, slander*, see also McKoon et al., 1993):

> These verbs share some properties with the *admire*-type psych-verbs [e.g. class 31.2 verbs]... While the *admire* verbs relate to a particular feeling that someone may have in reaction to something, these verbs relate to judgment or opinion that someone may have in reaction to something. (Levin, 1993, p. 196)

Thus, IC bias for these classes can be accounted for by the predicate decomposition schema already implemented in VerbNet, though considerable additional work is required to demonstrate that it *does* explain IC bias. For instance, while it is asserted that classes 31.1 (*frighten, surprise*), 30.2 (*discover, recognise*) and 33 (*praise, slander*) all contain the same IN_REACTION_TO component in their semantic structure, there does not appear to be any independent motivation for proposing that these verbs have this structure. In addition, VerbNet gives the same predicate decomposition to both classes 30.1 (*detect, hear*) and 30.2 (*discover, recognise*), while we find different IC biases for the two. Future research will need to address these issues.

Predicate decomposition and discourse structure

Both thematic role and predicate decomposition theories provide frameworks within which IC bias can be explained with reference to verbal semantics. Whether either will ultimately be sufficient is an open question (see also below). One reason to favour predicate decomposition at the outset is that the richer structures invoked may be more successful at accounting for the effects of different connectives. As noted above, pronoun resolution biases are a complex interaction of the verb and connective (Crinean & Garnham, 2006; Ehrlich, 1980; Kehler et al., 2008; Stewart et al., 1998):

(19) a. Sally_1 frightened Mary_2 because she_1...
b. Because Sally_1 frightened Mary_2 she_2...
c. Sally_1 frightened Mary_2, and then she_1

(20) a. Sally$_1$ feared Mary$_2$ because she$_2$...
b. Because Sally$_1$ feared Mary$_2$ she$_1$...
c. Sally$_1$ feared Mary$_2$, and then she$_1$

(21) a. Sally$_1$ criticized Mary$_2$ because she$_2$...
b. Because Sally$_1$ criticized Mary$_2$ she$_2$...
c. Sally$_1$ criticized Mary$_2$, and then she$_1$

Kehler et al. (2008) suggest that the different connectives set up different expectations about discourse continuations, and thus different aspects of the verb's semantics become relevant. Thus, in (19a), (20a), and (21a), the second clause explains the first and thus the pronoun should refer to the cause of the situation. In (19b), (20b), and (21b), the second clause refers to a consequence of the situation in the first, and thus the pronoun should refer to the affected entity to whom the consequence occurs. In (19c), (20c), and (21c), the two clauses describe a succession of events; in such cases, subjects tend to co-refer and verbal semantics is less relevant. Note that the pattern is different for the different verbs: while for *criticise*, the cause and affected entity are the same, for *frighten* and *fear*, they are different. While this can be described in terms of thematic roles (e.g. Stevenson, Crawley, & Kleinman, 1994), the fact that predicate decomposition straightforwardly allows the same argument to bear multiple roles with respect to the verb may account for these patterns more naturally.

CONTINUOUS VERSUS DISCRETE DISTRIBUTIONS OF BIAS

Several researchers have commented on the fact that IC bias appears to be continuously distributed, rather than bimodal (Garvey et al., 1974; Semin & Fiedler, 1991), and indeed in Experiment 1 we observed verbs at a wide range of IC biases. This fact is sometimes taken as evidence in favour of the *world knowledge* account, on which bias is necessarily graded, and against the *semantic structure* account, as early versions (e.g. Brown & Fish, 1983) predicted more categorical results.

A number of factors can mask underlying categoricity. First, categoricity may be masked by measurement error. Moreover, since the analyses above suggest that if IC bias is categorical, then there are at *least* three categories—subject biased, object biased, and nonbiased—and sampling from these three categories with some measurement error would give the appearance of a continuous distribution. This would be exacerbated if different verb *classes* have different strength biases. Thus, it may be that the underlying semantics of class 45.4 (*cool, improve*) is less causal that the semantics of class 31.1 (for instance, perhaps class 45.4 involves indirect causation, and class 31.1, direct causation). This would make the overall distribution look even more continuous.

Second, polysemous verbs may have one meaning that leads to one bias and another meaning that leads to another. Confusion over interpretation on the part of the participants would then lead to weaker biases in proportion to the confusion, further causing an apparent continuous distribution. Although we eliminated many polysemous verbs, we likely did not eliminate them all as polysemy itself is an open area of research. Similarly, some verbs were less natural in the sentential contexts employed, which could weaken participant intuitions. Finally, although IC bias is one factor in pronoun resolution, it is by no means the only (Ferstl et al., 2011; Kehler et al., 2008; Nappa & Arnold, 2009), and these other factors may affect different verbs differently, further smearing the distribution.

Thus, although a continuous distribution is often seen as evidence for the *world knowledge account*, in fact it is also consistent with the *semantic structure account*. However, the two theories account for such a distribution differently, and it remains for future work to tease these issues apart.

REVISITING WORLD KNOWLEDGE AND SEMANTIC FEATURES

While the *world knowledge* account and the *arbitrary semantic tag* account do not predict the effects of semantic structure, both accounts are in principle compatible with them. The semantic verb classes that we investigated are thought to reflect systematic differences in the conceptualisation of different types of events. Presumably the event concepts that underlie this semantic knowledge are also involved in representing our knowledge of typical (or specific) events in the world and their causes. Thus, on any hypothesis, we would expect to see a correlation between semantic verb classes and the contents of world knowledge, and so if one patterns with IC bias, the other will, too. Similarly, on the arbitrary tag hypothesis, the semantic tag of each verb must be learned, presumably on the basis of the utterances in which the verb appears. Since these utterances will describe specific events, as filtered through the human conceptual system, this information will be shaped by the same forces that have given rise to semantic verb classes and to our world knowledge.

However, the three proposals make quite different claims about the processes involved in IC biases and the means by which they are acquired. On the *semantic structures* hypothesis, the argument structure that is associated with a class of verbs encodes information that is relevant to inferring the causes of events. Consequently, IC can often be read off of the semantic representation of an utterance; once one has learned what a verb means, the IC bias comes for free. On the other two accounts, a verb's definition is insufficient. Thus, the most direct way to distinguish between these accounts would be to conduct studies in which participants are taught new verbs under conditions that controlled their knowledge of semantic structure, linguistic experience, and event knowledge.

With all this in mind, it is worth revisiting the most commonly cited evidence in favour of the *world knowledge* hypothesis: evidence that factors such as gender and social dominance relations affect IC judgments (Corrigan, 2001; LaFrance et al., 1997; Maass et al., 1989; Mannetti & De Grada, 1991). For instance, Corrigan (2001) found that participants were more likely to declare the object responsible in sentences like (22) than in sentences like (23), presumably because traitors are seen as more deserving of criticism than kings:

(22) The monarch criticized the traitor.
(23) The traitor criticized the monarch.

The *world knowledge* account explains that listeners access statistical knowledge about why monarchs and traitors criticise one another to ultimately decide who was the most likely cause. On the *semantic structure* and *arbitrary semantic tag* accounts, listeners initially determine who the cause was linguistically, but upon reflection, the listener can also access world knowledge about traitors and monarchs, and this may suggest additional hypotheses about what caused the event of criticism, modifying the original assessment.

Since it is possible to track moment-to-moment adjustments in listener's online assessments of pronoun resolution (Arnold, Brown-Schmidt, & Trueswell, 2007; Arnold, Eisenband, Brown-Schmidt, & Trueswell, 2000; Pyykkonen & Jarvikivi, 2010; Pyykkonen et al., 2010), pronoun resolution may be a particularly good place to look for distinctions in how different sources of information that influence IC bias are integrated online, though no such studies have been run.

CONCLUSION

This investigation of fine-grained semantic verb classes represents a promising direction in the study of IC and pronoun resolution in general. Previous investigations that have considered semantic verb classes have focused on coarse-grained distinctions, which typically resulted in only three or four verb classes (Rudolph & Forsterling, 1997). With few exceptions, the verb classes employed were based on intuitions about the causal structure encoded in the verb and thus were not independently motivated (but see Semin & Fiedler, 1988, 1991). The present study demonstrates that verb classification schemes based on syntactic patterns correlate with—and potentially explain—the direction of implicit causal biases.

These results suggest that we should be cautious in using IC to probe people's causal knowledge of world (Au, 1986; Corrigan & Stevenson, 1994; De Goede et al., 2009; Maass et al., 1989). To the extent that IC biases reflect knowledge of linguistic structure, they may provide a distorted picture of a person's nonlinguistic world knowledge. For example, measures of IC may overestimate children's understanding of causation (Au, 1986; Corrigan & Stevenson, 1994) if children are able to derive the semantic structures of verbs by tracking their syntactic properties (Gleitman, 1990) but do not fully understand the causal properties of the event types that they encode.

On a methodological level, this study considerably increases the number of English verbs for which IC biases have been reported (see Ferstl et al., 2011; Rudolph & Forsterling, 1997). By making these data publically available (see the appendixes), we hope to facilitate the creation of new experiments (see Goikoetxea et al., 2008, for a similar project in Spanish).

Finally, this discovery potentially provides a new probe into the semantic structures of verbs. The study of verb meaning within linguistics has focused largely on what constructions (e.g. dative, double object, progressive, etc.) a given verb can appear in. If these concerns are directly related to IC—and our analyses above suggest that they are—then implicit causal biases may provide a new data point to inform this project.

Manuscript received 1 October 2010
Revised manuscript received 11 March 2012
First published online 29 August 2012

REFERENCES

Ambridge, B., Pine, J. M., Rowland, C. F., & Young, C. R. (2008). The effect of verb semantic class and verb frequency (entrenchment) on children's and adults' graded judgements of argument-structure overgeneralization errors. *Cognition*, *106*, 87–129.

Arnold, J. E., Brown-Schmidt, S., & Trueswell, J. C. (2007). Children's use of gender and order-of-mention during pronoun comprehension. *Language and Cognitive Processes*, *22*, 527–565.

Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., & Trueswell, J. C. (2000). The rapid use of gender information: Evidence of the time course of pronoun resolution from eyetracking. *Cognition*, *76*, B13–B26.

Au, T. K. (1986). A verb is worth a thousand words: The causes and consequences of interpersonal events implicit in language. *Journal of Memory and Language*, *25*, 104–122.

Baker, M. C. (1988). *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.

Batchelder, W. H., Bershad, N. J., & Simpson, R. S. (1992). Dynamic paired-comparison scaling. *Journal of Mathematical Psychology*, *36*, 185–212.

Brown, R., & Fish, D. (1983a). The psychological causality implicit in language. *Cognition*, *14*, 237–273.

Brown, R., & Fish, D. (1983b). Are there universal schemas of psychological causality? *Archives de Psychologie*, *51*, 145–153.

Caramazza, A., Grober, E., Garvey, C., & Yates, J. (1977). Comprehension of anaphoric pronouns. *Journal of Verbal Learning and Verbal Behavior*, *16*, 601–609.

Corrigan, R. (2001). Implicit causality in language: Event participants and their interactions. *Journal of Language and Social Psychology*, *20*, 285–320.

Corrigan, R., & Stevenson, C. (1994). Children's causal attributions to states and events described by different classes of verbs. *Cognitive Development*, *9*, 235–256.

Crinean, M., & Garnham, A. (2006). Implicit causality, implicit consequentiality and thematic roles. *Language and Cognitive Processes*, *21*, 636–648.

Croft, W. A. (1991). *Syntactic categories and grammatical relations: The cognitive organization of information*. Chicago: University of Chicago Press.

Croft, W. A. (in press). *Verbs: Aspect and argument structure*. Oxford: Oxford University Press.

De Goede, D., Van Alphen, P., Mulder, E., Blokland, Y., Kerstholt, J., & Van Berkum, J. J. A. (2009). *The effect of mood on anticipation in language comprehension: An ERP study*. Presented at the 22nd Annual meeting of the CUNY Conference on Human Sentence Processing, Davis, CA.

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, *67*, 547–619.

Ehrlich, K. (1980). Comprehension of pronouns. *The Quarterly Journal of Experimental Psychology*, *32*, 247–255.

Elo, A. (1978). *The rating of chess players, past and present*. New York: Arco.

Featherstone, C., & Sturt, P. (2010). Because there was cause for concern: An investigation into the word-specific prediction account of the implicit causality effect. *Quarterly Journal of Experimental Psychology*, *63*, 3–15.

Ferstl, E. C., Garnham, A. & Manouilidou, C. (2011). Implicit causality biases in English: A corpus of 300 verbs. *Behavior Research Methods*, *43*, 124–135.

Forgas, J. P. (1995). Mood and judgment: The Affect Infusion Model (AIM). *Psychological Bulletin*, *117*, 39–66.

Frances, N., & Kucera, H. (1982). *Frequency analysis of English usage*. Boston, MA: Houghton Mifflin.

Frank, S. L., Koppen, M., Noordman, L. G. M., & Vonk, W. (2007). Coherence-driven resolution of referential ambiguity: A computational model. *Memory & Cognition*, *35*, 1307–1322.

Garnham, A., Traxler, M., Oakhill, J., & Gernsbacher, M. A. (1996). The locus of implicit causality effects in comprehension. *Journal of Memory and Language*, *35*, 517–543.

Garvey, C., & Caramazza, A. (1974). Implicit causality in verbs. *Linguistic Inquiry*, *5*, 459–464.

Garvey, C., Caramazza, A., & Yates, J. (1974). Factors influencing assignment of pronoun antecedents. *Cognition*, *3*, 227–243.

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition*, *1*, 3–55.

Goikoetxea, E., Pascual, G., & Acha, J. (2008). Normative study of the implicit causality of 100 interpersonal verbs in Spanish. *Behavior Research Methods*, *40*, 760–772.

Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago, IL: University of Chicago Press.

Greene, S. B., & McKoon, G. (1995). Telling something we can't know: Experimental approaches to verbs exhibiting implicit causality. *Psychological Science*, *6*, 262–270.

Guerry, M., Gimenes, M., Caplan, D., & Rigalleau, F. (2006). How long does it take to find a cause? An online investigation of implicit causality in sentence production. *The Quarterly Journal of Experimental Psychology*, *59*, 1535–1555.

Harley, H. (1995). *Subjects, events, and licensing* (Doctoral dissertation, MIT, Cambridge, MA).

Hartigan, J. A., & Hartigan, P. M. (1985). The dip test of unimodality. *The Annals of Statistics*, *13*, 70–84.

Hartshorne, J. K. (submitted). Pronoun interpretation, causal attribution, and event participants. Manuscript submitted for publication.

Hoffman, C., & Tchir, M. A. (1990). Interpersonal verbs and dispositional adjectives: The psychology of causality embodied in language. *Journal of personality and social psychology*, *58*, 765–778.

Jackendoff, R. (1990). *Semantic structures*. Cambridge, MA: The MIT Press.

Jameson, K. A., Appleby, M. C., & Freeman, L. C. (1999). Finding an appropriate order for a hierarchy based on probabilistic dominance. *Animal Behavior*, *57*, 991–998.

Kehler, A. (2002). *Coherence, reference, and the theory of grammar*. Stanford, CA: CSLI Publications.

Kehler, A., Kertz, L., Rohde, H., & Elman, J. L. (2008). Coherence and coreference revisited. *Journal of Semantics*, *25*, 1–44.

Kipper, K., Korhonen, A., Ryant, N., & Palmer, M. (2008). A large-scale classification of English verbs. *Language Resources and Evaluation Journal*, *42*, 21–40.

Koornneef, A. W., & Van Berkum, J. J. A. (2006). On the use of verb-based implicit causality in sentence comprehension: Evidence from self-paced reading and eye tracking. *Journal of Memory and Language*, *54*, 445–465.

Korhonen, A., & Briscoe, T. (2004). *Extended lexical-semantic classification of English verbs*. Proceedings of the HLT/NAACL workshop on computational lexical semantics, Boston, MA.

Korhonen, A., & Ryant, N. (2005). *53 novel lexical-semantic verb classes*. Unpublished manuscript.

LaFrance, M., Brownell, H., & Hahn, E. (1997). Interpersonal verbs, gender, and implicit causality. *Social Psychology Quarterly*, *60*, 138–152.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. Chicago, IL: University of Chicago Press.

Levin, B., & Rappaport Hovav, M. (2005). *Argument realization.* Research Surveys in Linguistics Series. Cambridge: Cambridge University Press.

Lidz, J., Gleitman, H., & Gleitman, L. (2003). Understanding how input matters: Verb learning and the footprint of universal grammar. *Cognition*, *87*, 151–178.

Long, D. L., & De Ley, L. (2000). Implicit causality and discourse focus: The interaction of text and reader characteristics in pronoun resolution. *Journal of Memory and Language*, *42*, 545–570.

Maechler, M. (2009). *Diptest: Hartigan's dip test statistic for unimodality – corrected code*. Retrieved from http://cran.r-project.org/web/packages/diptest/index.html

Mannetti, L., & De Grada, E. (1991). Interpersonal verbs: Implicit causality of action verbs and contextual factors. *European Journal of Social Psychology*, *21*, 429–443.

McDonald, J. L., & MacWhinney, B. (1995). The time course of anaphor resolution: Effects of implicit verb causality and gender. *Journal of Memory and Language*, *34*, 543–566.

McKoon, G., Greene, S. B., & Ratcliff, R. (1993). Discourse models, pronoun resolution, and the implicit causality of verbs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 1040–1052.

Moens, M., & Steedman, M. (1988). Temporal ontology and temporal reference. *Computational Linguistics*, *14*, 15–28.

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, *17*(2), 357–374.

Nappa, R., & Arnold, J. (2009). *Interactions between cue informativeness and reliability in the resolution of referential ambiguity*. Presented at the 50th Annual Meeting of the Psychonomic Society, Boston, MA.

Pesetsky, D. (1995). *Zero syntax: Experiencers and cascades*. Cambridge, MA: The MIT Press.

Pickering, M. J., & Majid, A. (2007). What are implicit causality and consequentiality? *Language and Cognitive Processes*, *22*, 780–788.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: The MIT Press.

Pustejovsky, J. (1995). *The generative lexicon*. Cambridge, MA: The MIT Press.

Pyykkonen, P., & Jarvikivi, J. (2010). Activation and persistence of implicit causality information in spoken language comprehension. *Experimental Psychology*, *57*, 5–16.

Pyykkonen, P., Matthews, D., & Jarvikivi, J. (2010). Three-year-olds are sensitive to semantic prominence during online language comprehension: A visual world study of pronoun resolution. *Language and Cognitive Processes*, *25*, 115–129.

R Development Core Team. (2009). *R: A language and environment for statistical computing*. Retrieved from http://www.R-project.org

Rappaport Hovav, M., & Levin, B. (1998). Building verb meanings. In M. Butt & W. Geuder (Eds.), *The projection of arguments: Lexical and compositional factors* (pp. 97–134). Stanford: Center for the Study of Language and Information.

Rudolph, U., & Forsterling, F. (1997). The psychological causality implicit in verbs: A review. *Psychological Bulletin*, *121*, 192–218.

Semin, G. R., & Fiedler, K. (1988). The cognitive functions of linguistic categories in describing persons: Social cognition and language. *Journal of Personality and Social Psychology*, *54*, 558–568.

Semin, G. R., & Fiedler, K. (1991). The linguistic category model, its bases, applications and range. *European Review of Social Psychology*, *2*, 1–30.

Semin, G. R., & Marsman, J. G. (1994). "Multiple inference-inviting properties" of interpersonal verbs: Event instigation, dispositional inference, and implicit causality. *Journal of Personality and Social Psychology*, *67*, 836–849.

Shen, M., & Yang, Y. (2006). The effects of implicit verb causality and accentuation on pronoun processing. *Acta Psychologica Sinica*, *38*(4), 497–506.

Stevenson, R. J., Crawley, R. A., & Kleinman, D. (1994). Thematic roles, focus and the representation of events. *Language and Cognitive Processes*, *9*, 519–548.

Stewart, A. J., Pickering, M. J., & Sanford, A. J. (1998). Implicit consequentiality. In *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (pp. 1031–1036). Mahwah, NJ: Lawrence Erlbaum Associates.

Stewart, A. J., Pickering, M. J., & Sanford, A. J. (2000). The time course of the influence of implicit causality information: Focusing versus integration accounts. *Journal of Memory and Language*, *42*, 423–443.

Talmy, L. (2000a). *Toward a cognitive semantics: Vol. 1: Concept structuring systems*. Cambridge, MA: The MIT Press.

Talmy, L. (2000b). *Toward a cognitive semantics: Typology and processing in concept structuring*. Cambridge, MA: The MIT Press.

Van Valin, R. D. (1993). A synopsis of role and reference grammar. In R. D. Van Valin (Ed.), *Advances in role and reference grammar* (pp. 1–164). Amsterdam: John Benajamins.

Van Valin, R. D. (2004). Semantic macroroles in role and reference grammar. In R. Kailuweit & M. Hummel (Eds.), *Semantische rollen* (pp. 62–82). Tubingen: Gunter Narr Verlag.

Whorf, B. L. (1956). *Language, thought and reality: Selected writings of Benjamin Lee Whorf* (J. B. Carroll (Ed.).). Cambridge, MA: The MIT Press.

APPENDIX 1
Experiment 1 stimuli and results

Verb N Object-Bias
abandons 49 67%
abolishes 47 47%
accelerates 52 46%
accepts 49 57%
accommodates 46 70%
accompanies 34 74%
accomplishes 38 29%
accuses 50 86%
achieves 56 41%
acknowledges 32 72%
acquires 40 63%
adapts 43 63%
addresses 51 80%
adjusts 48 56%
admires 51 82%
admits 51 53%
adopts 47 79%
advances 40 58%
advises 47 38%
advocates 48 56%
affects 34 24%
affirms 53 79%
aids 45 53%
alerts 47 38%
alters 51 53%
analyzes 48 65%
announces 47 55%
answers 41 54%
anticipates 37 51%
applauds 38 61%
applies 44 66%
appoints 41 80%
appraises 57 35%
appreciates 46 80%
approaches 45 76%
approves 58 53%
argues 33 48%
arouses 46 33%
arranges 44 52%
arrests 51 82%
ascertains 52 50%
asks 56 75%
assembles 46 41%
asserts 46 43%
assesses 64 61%
assigns 43 91%
assists 56 63%
assures 39 36%
attaches 51 71%
attacks 58 71%
attains 45 49%
attempts 35 49%
attracts 54 20%
authorizes 48 52%
avoids 46 85%
awaits 39 51%
backs 44 55%
balances 45 44%
banishes 51 82%
bathes 65 75%
bears 49 45%
beats 39 56%
begs 42 55%
beholds 40 78%
believes 56 68%
bends 54 39%
benefits 45 24%
betrays 48 35%
betters 52 23%
bites 37 51%
blames 45 80%
blends 47 53%
blesses 42 60%
blocks 58 64%
boasts 43 37%
boils 31 71%
boosts 54 61%
bores 47 21%
borrows 45 78%
bothers 62 48%
brushes 41 56%
builds 56 48%
bumps 46 61%
buries 41 71%
burns 50 66%
buys 51 78%
calculates 39 44%
calls 47 60%
calms 44 55%
cancels 49 51%
captures 47 79%
carries 64 61%
casts 49 84%
catches 55 53%
celebrates 51 71%
certifies 49 69%
challenges 46 61%
changes 52 56%
characterizes 41 51%

charges 50 84%
chases 43 65%
checks 44 70%
cherishes 50 78%
chokes 42 86%
chooses 42 90%
cites 49 67%
claims 45 80%
clarifies 53 55%
classifies 53 70%
cleans 56 54%
clears 42 57%
climbs 46 46%
closes 46 63%
collects 44 50%
combats 41 63%
commands 51 51%
commends 54 70%
communicates 52 52%
compares 58 69%
compels 51 43%
composes 56 57%
comprehends 46 46%
computes 53 47%
conceals 29 62%
concedes 42 52%
concerns 60 42%
concludes 48 52%
condemns 52 88%
conducts 63 52%
confesses 53 70%
confirms 45 40%
confronts 54 78%
confuses 50 38%
congratulates 44 70%
conquers 45 53%
considers 39 74%
constructs 42 48%
consults 43 84%
contacts 54 56%
contemplates 58 78%
contracts 42 60%
contradicts 47 51%
contrasts 50 52%
contributes 47 40%
controls 53 26%
converts 51 41%
conveys 56 68%
convinces 47 43%
cooks 46 63%
cools 51 45%
coordinates 51 53%
corrects 47 51%
counteracts 45 49%
covers 49 55%
cracks 48 58%
creates 53 28%
criticizes 47 83%
crosses 56 57%
cures 54 39%
curses 49 82%
curtails 45 64%
cuts 45 73%
damages 49 73%
dances 40 48%
dares 40 55%
dates 52 77%
decides 53 51%
declares 47 49%
declines 49 78%
defeats 48 40%
defends 46 70%
defies 44 34%
defines 45 60%
delays 40 48%
delegates 40 60%
demands 59 37%
demonstrates 64 53%
denies 51 69%
denounces 43 86%
describes 52 77%
deserves 39 36%
designs 52 42%
desires 44 89%
despises 45 84%
destroys 52 75%
detects 63 59%
devises 41 49%
devotes 44 55%
digs 50 38%
directs 50 56%
discerns 48 42%
discloses 47 64%
discourages 39 59%
discovers 48 58%
discusses 54 57%
dislikes 55 93%
dismisses 40 93%
displays 43 79%
disputes 40 55%
disrupts 61 52%
dissolves 54 46%
distinguishes 54 76%
distorts 49 43%
distributes 35 60%
disturbs 45 36%
divides 33 64%
divorces 54 80%
dodges 42 86%
dominates 58 38%
doubles 52 62%
doubts 45 80%
drags 49 59%
drains 54 48%
draws 52 54%
dresses 48 63%
dries 40 65%
drills 35 51%
drinks 48 44%
drops 35 66%
earns 54 43%
eats 50 60%
educates 44 55%
elaborates 51 65%
elects 52 83%
eliminates 41 76%
embraces 49 57%
emphasizes 53 60%
employs 57 77%
enables 44 39%
enacts 36 47%
encounters 46 48%
encourages 50 54%
endorses 44 82%
ends 37 59%
endures 47 66%
engages 34 62%
enhances 48 46%
enjoys 50 86%
enlarges 44 43%
enlists 54 89%
enriches 44 27%
enrolls 60 73%
ensures 51 55%
entertains 43 35%
equals 52 46%
erects 44 48%
escapes 58 41%
escorts 45 56%
establishes 42 50%
estimates 35 37%
evaluates 44 59%
exaggerates 54 30%
examines 47 70%
exceeds 51 18%
excludes 46 80%
excuses 35 83%
executes 53 74%
exhibits 34 76%
expects 47 55%
explains 39 72%
explodes 55 44%
exploits 53 68%
explores 43 63%
exposes 56 73%
expresses 50 60%
extracts 48 46%
faces 42 43%
facilitates 46 43%
fails 50 42%
fans 44 57%
favors 41 80%
fears 49 80%
features 52 81%
feeds 41 78%
fetches 45 64%
fights 55 67%
files 42 40%
finances 49 55%
finds 49 51%
finishes 58 48%
fires 45 82%
fits 46 48%
fixes 56 46%
flashes 54 41%
follows 46 59%
fools 62 55%
forbids 40 85%
forces 55 44%
forestalls 51 65%
forgets 51 55%
forgives 51 55%
formulates 45 38%
frames 44 55%
frees 40 60%
freezes 47 66%
frightens 58 24%
frustrates 50 28%
fulfills 45 20%
furnishes 55 47%

gathers 45 44%
generates 51 55%
gives 44 66%
governs 56 52%
grabs 42 69%
grasps 41 56%
greets 44 57%
groups 54 44%
guards 41 54%
guesses 50 40%
guides 42 62%
handles 44 61%
hands 37 54%
hangs 60 72%
hates 54 87%
hears 47 45%
heats 43 67%
heeds 38 68%
helps 38 50%
hides 44 86%
hires 51 82%
hits 67 72%
holds 51 67%
honors 41 78%
houses 50 62%
hunts 37 62%
hurries 41 49%
hurts 58 76%
identifies 45 87%
ignores 55 82%
illustrates 44 75%
imagines 52 37%
imitates 40 40%
imparts 30 63%
implies 36 44%
imposes 40 45%
improves 46 35%
includes 55 82%
indulges 47 77%
inflicts 51 73%
influences 50 42%
informs 44 55%
inherits 45 31%
inhibits 52 44%
initiates 43 53%
injects 44 82%
inspects 41 73%
installs 40 53%
intends 57 67%
intensifies 41 59%
interprets 49 65%
interrupts 41 41%
introduces 46 59%
invents 44 48%
investigates 48 69%
invites 52 79%
involves 45 78%
isolates 39 72%
issues 59 75%
jeopardizes 48 44%
joins 52 71%
judges 54 63%
justifies 35 49%
keeps 48 79%
kicks 41 66%
kills 43 84%
kisses 69 68%
knocks 43 79%
knows 49 63%
lashes 50 88%
launches 36 72%
leads 48 50%
learns 37 78%
leaves 47 85%
liberates 38 53%
lifts 44 59%
lights 50 62%
likes 50 86%
limits 45 69%
lines 44 52%
links 28 82%
lists 46 50%
locates 43 30%
loses 41 51%
loves 49 76%
lowers 42 62%
mails 53 57%
manages 58 52%
manipulates 48 40%
marks 43 79%
marries 59 75%
masters 41 29%
matches 45 42%
measures 43 58%
meets 42 81%
melts 57 46%
mentions 44 82%
merges 44 52%
minds 39 69%
minimizes 47 68%
misses 59 58%
mistakes 49 37%
mixes 42 40%
modifies 48 60%
monopolizes 50 56%
moves 41 85%
murders 49 80%
names 46 54%
nears 47 66%
needs 56 63%
neglects 56 68%
negotiates 49 55%
notes 42 57%
notices 47 74%
notifies 39 56%
obeys 44 50%
obscures 59 56%
observes 40 68%
obstructs 44 61%
obtains 50 76%
offends 48 38%
offers 44 61%
opens 51 57%
opposes 41 71%
orders 45 58%
organizes 41 56%
outgrows 41 49%
outlines 40 50%
overcomes 53 26%
overlooks 44 82%
owes 51 43%
owns 41 59%
packs 45 58%
paints 49 47%
pardons 36 72%
passes 47 68%
pauses 46 61%
pays 36 78%
performs 53 45%
perpetuates 48 42%
persuades 52 40%
phones 57 61%
picks 49 94%
plans 45 71%
plants 57 51%
pleads 49 63%
pleases 46 35%
plots 69 51%
polishes 49 37%
portrays 40 48%
poses 43 65%
positions 46 61%
possesses 51 45%
postpones 48 56%
pours 55 62%
practices 47 38%
praises 37 68%
preaches 52 44%
precludes 47 72%
predicts 47 40%
prefers 47 87%
prepares 68 51%
prescribes 58 57%
presents 50 66%
preserves 53 66%
presses 42 71%
pretends 44 43%
prints 48 50%
probes 49 71%
proclaims 45 60%
procures 41 54%
produces 57 42%
professes 57 49%
programs 48 42%
promises 60 52%
promotes 67 79%
propels 43 67%
proposes 41 73%
protects 43 81%
protests 39 82%
proves 41 24%
pulls 54 63%
purchases 38 74%
pursues 57 77%
pushes 43 70%
puts 40 45%
puzzles 44 30%
questions 47 72%
quits 46 50%
quotes 43 72%
races 56 45%
rates 48 63%
rationalizes 49 53%
re-examines 47 74%
reaches 35 46%
reads 51 67%
rebuilds 38 39%
rebuts 46 54%
recalls 48 75%
receives 38 26%

recognizes 29 79%
recommends 43 81%
reconciles 35 57%
reconsiders 52 69%
reconstructs 53 40%
records 44 48%
recovers 51 31%
refuses 55 78%
regards 43 72%
registers 47 64%
reinforces 46 61%
rejects 45 82%
relates 46 63%
releases 57 82%
relieves 50 52%
relinquishes 46 83%
remembers 58 76%
reminds 42 69%
renders 44 45%
renews 43 60%
rents 51 61%
repairs 59 47%
repays 35 40%
repeats 45 64%
replaces 52 60%
replenishes 43 42%
replies 45 51%
reports 56 45%
represents 56 55%
reproduces 50 52%
requests 42 69%
requires 46 48%
rescues 90 46%
resembles 51 31%
resents 48 90%
reserves 50 70%
resists 47 66%
resolves 51 37%
respects 53 85%
restores 40 53%
restrains 44 70%
restricts 56 79%
resumes 48 52%
retains 54 83%
returns 29 34%
reveals 41 71%
reverses 45 60%
reviews 44 61%
revises 38 50%
revives 48 40%
rings 65 62%
rips 34 68%
risks 48 44%
rolls 44 77%
rubs 49 57%
ruins 52 60%
rules 44 32%
sacrifices 48 69%
sails 37 46%
sanctions 59 78%
satisfies 53 32%
saves 51 63%
saws 44 64%
says 53 74%
schools 56 68%
screens 58 43%
scrubs 40 68%
searches 44 82%
secures 47 55%
seeks 43 77%
sees 58 64%
seizes 46 80%
selects 56 89%
sells 51 76%
senses 45 44%
separates 43 56%
serves 42 50%
services 45 49%
sews 42 48%
shakes 50 74%
shapes 35 49%
shares 52 52%
shaves 54 67%
shoots 60 90%
shortens 65 55%
shouts 53 66%
shows 50 64%
shrinks 57 60%
shuts 36 67%
signals 60 52%
signs 56 46%
simplifies 44 77%
simulates 56 55%
sings 48 38%
sinks 47 47%
skips 40 90%
slides 41 59%
slows 45 56%
smacks 52 73%
smells 35 71%
smokes 41 39%
smooths 30 23%
snaps 49 59%
snatches 48 69%
soaks 46 67%
softens 54 41%
sorts 43 37%
sovles 58 40%
spares 38 74%
specifies 48 63%
spells 43 44%
spins 47 49%
sponsors 45 80%
spots 51 78%
squeezes 54 80%
stamps 44 68%
states 45 56%
steals 39 59%
stimulates 57 33%
stirs 47 45%
stops 42 71%
stores 54 41%
straightens 39 64%
strengthens 37 24%
stresses 47 34%
stretches 41 54%
strikes 43 86%
strips 41 73%
studies 50 76%
submits 41 59%
subsidizes 46 74%
substitutes 60 55%
succeeds 51 31%
sues 44 68%
suffers 48 38%
suggests 34 38%
suits 44 43%
supervises 50 56%
supplies 70 54%
supports 54 76%
suppresses 36 72%
surprises 62 42%
surrenders 37 68%
surrounds 52 63%
survives 49 27%
suspects 50 82%
sustains 53 57%
swears 52 60%
sweeps 50 32%
switches 48 60%
symbolizes 64 30%
tackles 58 84%
taps 46 65%
tastes 43 44%
taxes 29 72%
teaches 50 52%
teases 52 79%
tells 56 70%
terminates 44 84%
testifies 47 51%
tests 46 67%
thanks 43 53%
threatens 38 71%
throws 51 59%
ties 48 67%
toasts 40 75%
tolerates 50 66%
tosses 41 83%
touches 46 61%
traces 49 71%
trades 44 66%
trains 65 68%
transfers 33 85%
transforms 51 39%
translates 49 53%
treats 51 51%
trims 47 62%
troubles 42 50%
trusts 47 89%
twists 41 66%
uncovers 47 64%
underestimates 51 65%
undergoes 41 54%
undermines 37 41%
understands 50 40%
undertakes 47 53%
unites 47 40%
unloads 46 65%
upholds 47 51%
urges 33 52%
uses 45 64%
utilizes 51 67%
verifies 50 52%
views 42 71%
violates 45 53%
visits 44 68%
volunteers 48 88%
wakes 54 63%
wants 54 85%
warns 49 53%

washes 52 71%
wastes 53 36%
watches 45 71%
weakens 43 53%
wears 53 38%
weighs 45 69%
welcomes 42 76%
whips 59 80%
whispers 47 51%
widens 50 66%
winds 48 65%
wins 47 45%
wipes 37 70%
wishes 50 56%
withdraws 49 65%
witnesses 61 38%
worries 43 21%
worships 51 75%
wraps 43 56%
writes 63 51%
yields 38 61%

APPENDIX 2
Experiment 2 stimuli and results

Verb N Object-Bias
abashes 62 55%
abhors 277 82%
admires 268 89%
adores 302 88%
affects 62 24%
afflicts 59 37%
affronts 58 36%
aggravates 60 27%
agitates 46 28%
agonizes 48 35%
alarms 58 34%
alienates 69 78%
amazes 53 23%
amuses 60 18%
angers 66 35%
annoys 55 27%
antagonizes 62 58%
appalls 66 53%
appeases 53 36%
appreciates 258 79%
arouses 68 19%
assuages 49 53%
astonishes 62 23%
astounds 52 21%
awes 64 39%
baffles 60 27%
beguiles 58 41%
bewilders 49 35%
bewitches 53 25%
boggles 60 30%
bores 54 24%
bothers 54 31%
bugs 46 24%
calms 58 24%
captivates 64 33%
chagrins 60 45%
charms 61 20%
cheers 60 53%
cherishes 267 87%
chills 54 30%
comforts 57 46%
concerns 48 42%
confounds 55 29%
confuses 70 27%
consoles 47 49%
contents 56 46%
convinces 50 26%
cows 59 63%
crushes 52 52%
cuts 60 60%
daunts 72 32%
dazes 48 29%
dazzles 63 22%
dejects 59 61%
delights 57 19%
demolishes 54 57%
demoralizes 56 50%
deplores 287 85%
depresses 64 33%
despises 269 89%
detests 267 86%
devastates 49 22%
disappoints 58 16%
disarms 62 40%
discombobulates 54 35%
discomfits 51 37%
discomposes 42 48%
disconcerts 45 29%
discourages 50 64%
disdains 278 85%
disgraces 53 40%
disgruntles 55 36%
disgusts 57 30%
disheartens 60 30%
disillusions 54 33%
dislikes 269 89%
dismays 50 36%
dispirits 62 44%
displeases 50 24%
disquiets 54 46%
dissatisfies 44 36%
distracts 45 24%
distresses 55 42%
distrusts 291 82%
disturbs 64 22%
dreads 287 86%
dumbfounds 64 27%
elates 60 38%
electrifies 52 21%
embarrasses 58 29%
emboldens 52 27%
enchants 49 27%
encourages 43 40%
engages 55 51%
engrosses 60 32%
enjoys 268 87%
enlightens 41 20%
enlivens 61 33%
enrages 63 17%
enraptures 54 35%
entertains 57 42%
enthralls 48 27%
enthuses 57 37%
entices 32 28%
entrances 48 21%
envies 272 86%
esteems 267 76%
exalts 306 78%
exasperates 59 24%
excites 60 20%
execrates 261 70%
exhausts 59 36%
exhilarates 58 31%
fancies 296 84%
fascinates 69 26%
favors 296 83%
fazes 53 42%
fears 291 85%
flabbergasts 53 26%
flatters 50 32%
floors 46 37%
flusters 58 31%
frightens 59 19%
frustrates 45 22%
galls 56 38%
galvanizes 49 53%
gladdens 47 17%
gratifies 58 41%
grieves 51 49%
harasses 55 73%
hates 270 85%
haunts 58 31%
heartens 50 36%
horrifies 57 25%
humbles 49 49%
humiliates 59 58%
hurts 59 71%
hypnotizes 50 44%

idolizes 289 87%
impresses 61 15%
incenses 42 38%
infuriates 64 23%
inspires 54 19%
insults 62 68%
interests 58 36%
intimidates 48 17%
intoxicates 56 29%
intrigues 57 30%
invigorates 64 16%
irks 62 27%
irritates 55 22%
jars 56 46%
jollifies 56 39%
jolts 49 47%
laments 289 75%
likes 270 87%
loathes 278 86%
loves 270 86%
lulls 51 39%
maddens 51 33%
mesmerizes 56 18%
miffs 48 19%
misses 263 68%
mollifies 58 57%
mortifies 70 27%
mourns 284 67%
moves 62 53%
muddles 61 36%
mystifies 64 25%
nauseates 65 31%
nettles 48 40%
numbs 60 45%
obsesses 59 42%
offends 53 23%
outrages 56 36%
overawes 69 41%
overwhelms 62 18%
pacifies 57 42%
pains 59 34%
peeves 47 32%
perplexes 53 23%
perturbs 70 21%
piques 53 47%
pities 288 84%
placates 63 59%
plagues 67 28%
pleases 58 22%
preoccupies 56 21%
prizes 286 85%
provokes 64 41%
puzzles 59 34%
rankles 52 42%
reassures 47 36%
refreshes 50 28%
regrets 272 72%
relaxes 44 25%
relieves 62 50%
relishes 292 85%
repels 60 38%
repulses 59 31%
resents 267 85%
respects 277 84%
reveres 262 84%
revitalizes 50 22%
revolts 66 36%
riles 66 32%
ruffles 49 47%
saddens 63 22%
satisfies 64 23%
savors 304 80%
scandalizes 64 48%
scares 70 13%
shakes 43 88%
shames 47 53%
shocks 47 26%
sickens 49 35%
sobers 68 31%
solaces 77 44%
soothes 53 36%
spellbinds 57 19%
spooks 55 22%
staggers 50 32%
stands 307 64%
startles 62 26%
stimulates 53 38%
stings 54 65%
stirs 65 52%
strikes 60 78%
stumps 57 44%
stuns 65 25%
stupefies 58 29%
supports 276 63%
surprises 57 28%
tantalizes 54 22%
teases 60 80%
tempts 49 16%
terrifies 57 26%
terrorizes 56 48%
threatens 64 58%
thrills 64 25%
throws 45 64%
tickles 54 52%
tires 58 22%
titillates 56 25%
tolerates 282 63%
torments 44 59%
touches 54 61%
transports 63 65%
treasures 280 86%
tries 59 49%
troubles 56 32%
trusts 289 85%
unnerves 38 18%
unsettles 50 20%
uplifts 49 22%
upsets 47 38%
values 272 86%
venerates 268 76%
vexes 63 37%
wearies 48 21%
worries 62 26%
worships 303 84%
wounds 47 60%
wows 53 26%

APPENDIX 3
Re-analysis of Ferstl et al. (2011)

In order to test the reliability of our findings, we reanalysed the 305 verbs reported in Ferstl et al. (2011). Participants in that study completed sentence fragments (*Sally* frightened *John because*. . .), and IC bias was calculated based on who the continuations referred to (Sally or John), which was typically unambiguous because of the use of a gendered pronoun (*he* or *she*). The authors first selected 109 verbs from previous studies (Au, 1986; Crinean & Garnham, 2006; Rudolph, 2008). They classified these verbs according to Levin's (1993) verb classes (a precursor to VerbNet) and chose additional verbs from those classes.

Results and Discussion

Results were re-analysed following the analyses in Experiment 1. We identified 211 monosemic verbs with a mean object bias of 49.3%, which was used as the baseline for subsequent analyses (see Experiment 1). We first conducted analyses according to the Brown and Fish taxonomy, the Rudolph and Forsterling/McKoon et al. taxonomy, the Au taxonomy, and the Linguistic Category Model (Table A1). As predicted, experiencer-stimulus verbs were object-biased and stimulus-experiencer verbs were subject biased. Brown & Fish's action verbs are significantly object biased, contrasting both with the theory's prediction and the results of Experiment 1. Au's action-agent verbs and Rudolph and Forsterling/McKoon et al.'s agent-patient verbs are nonbiased, rather than subject biased as predicted. The first author coded the Ferstl et al. (2011)

verbs both according to the broad interpretation and the narrow interpretation of the Linguistic Category Model's descriptive action verbs (see Experiment 1 method). On the narrow interpretation, descriptive action verbs were non-biased ($t < 1$) and interpretive action verbs were significantly object biased ($t = 4.88$, $p < .00001$). On the broad interpretation, descriptive action verbs—which are predicted to be nonbiased or weakly subject biased—were significantly object biased, whereas interpretive action verbs—predicted to be subject biased—were nonbiased.

We then reanalysed this data according to the VerbNet verb classes. We found 92 verbs in class 31.1 (*frighten, surprise*), 33 in class 31.2 (*fear, love*), 48 in class 33 (*praise, slander*), 7 in class 36.2 (*court, cuddle*), and no more than 4 in each of 20 additional classes. As in Experiment 1, results were fit stepwise bi-directionally to a linear model with these classes as predictors. Classes 31.1 (*frighten, surprise*), 31.2 and 33 (*praise, slander*) emerged as significant predictors. Class 31.1 was significantly subject biased, and classes 31.2 and 33 (*praise, slander*) were significantly object biased (Table A2), replicating results from Experiment 1. Class 36.2 (*court, cuddle*), which consists largely of verbs of courtship, showed no significant bias ($M_{\text{diff}} = -8\%$, $SD = 18\%$, $t = 1.10$, $p = .31$).

These results explain why the two versions of the Linguistic Category Model produced different results. Most class 33 verbs (*praise, slander*) involve communication, which on the narrow interpretation were classified as Interpretive Action verbs (communication may proceed in oral or written form and thus has no physically invariant component), and which on the broad interpretation were classified as Descriptive Action Verbs (following the examples given by the authors, which include communication verbs like *summon*). Thus, whichever class included class 33 verbs (*praise, slander*) was object biased. Thus the descriptive action/interpretive action distinction appears to be doing no work beyond what is done by identifying class 33 (*praise, slander*).

Thus, the reanalysis of Ferstl et al. (2011) confirms several important conclusions from Experiments 1 and 2. First, the finding of strong consistent biases for three VerbNet classes—31.1 (*frighten, surprise*), 31.2 (*fear, love*), and 33 (*praise, slander*)—was replicated. Second, while previous taxonomies roughly characterise psych verbs correctly, they all made incorrect predictions about action verbs. Moreover, the Brown and Fish taxonomy, the Au taxonomy, and the Linguistic Category Model taxonomy all saw the direction of bias for at least one class change relative to Experiment 1. This does not appear to be random variation but due to the fact that the biases of verbs vary systematically *within* verb classes used by these taxonomies. Thus, Au's action-patient verbs were more object biased in the Ferstl et al. data relative to Experiment 1 because the

TABLE A1

Results from Ferstl et al. (in press) for the three previous semantic structure accounts by verb class, with class object bias mean (standard deviation), compared to grand average for monosemic verbs. Note that all three employ the same experiencer-stimulus and stimulus-experiencer classes

		Object bias	
Class	N	*Diff. from Mean*	*Significance*
Brown & Fish, Rudolph & Forsterling, and Linguistic Category Model			
Experiencer-stimulus	34	+35% (11%)	$t = 18.56$, $p < .00001$
Stimulus-experiencer	94	−22% (19%)	$t = 11.26$ $p < .00001$
Brown and Fish			
Agent-patient	83	+11% (21%)	$t = 4.57$, $p = .00001$
Rudolph & Forstering/McKoon et al.			
Agent-evocator	48	+17% (12%)	$t = 6.59$, $p < .00001$
Agent-patient	35	+3% (24%)	$t < 1$
Au			
Action-agent	50	+16% (23%)	$t = 6.39$, $p < .00001$
Action-patient	33	+2% (18%)	$t < 1$
Linguistic Category Model			
Narrow			
Descriptive action	10	+2% (25%)	$t < 1$
Interpretive action	73	+12% (21%)	$t = 4.88$, $p = .000006$
Broad			
Descriptive action	41	+16% (19%)	$t = 5.36$, $p < .00001$
Interpretive action	42	+5% (22%)	$t = 1.59$, $p = .12$

former included a larger proportion of class 33 verbs (*praise, slander*). Brown and Fish's action verb class changed from nonbiased (Experiment 1) to object biased (Ferstl et al., 2011) for the same reason. Similar considerations apply to the Linguistic Category Model. While the Rudolph and Forsterling taxonomy fares somewhat better, it is still unable to distinguish subject biased and nonbiased action verbs, and if there are any additional classes of object-biased action verbs, it would be unable to identify those as well. Thus, the value of utilising more narrowly-defined verb classes such as those in VerbNet is confirmed.

TABLE A2

VerbNet verb classes for which at least five verbs were tested, with class object bias mean (standard deviation), compared to grand average for monosemic verbs. Data taken from Ferstl et al. (in press)

		Object bias		
Class	N	*Diff. From Mean*	*Significance*	*Examples*
31.1	92	−22% (20%)	$t = 11.04$ $p < .00001$	calms, confuses, frustrates, troubles
31.2	33	+35% (10%)	$t = 18.16$, $p < .00001$	admires, cherishes, despises, loves
33	48	+17% (18%)	$t = 6.60$, $p < .00001$	blames, congratulates, thanks
36.2	7	−8% (18%)	$t = 1.10$, $p = .31$	courted, cuddled, divorced