WHAT IS IMPLICIT CAUSALITY: WORLD KNOWLEDGE, AN ARBITRARY FEATURE OR AN EFFECT OF SEMANTIC STRUCTURE?

Joshua K. Hartshorne
Jesse Snedeker
Harvard University

Running title: What is implicit causality?

Word count: 6,909 + references (1,294) + appendix (2,958) + abstract (216)

Send Correspondence to:
Joshua Hartshorne
33 Kirkland Street
WJH 1120
Cambridge, MA 02138
Email: jharts@wjh.harvard.edu
Tel: 617-496-4486
Fax: 617-495-3728

Acknowledgements:

The authors wish to thank Alfonso Caramazza, Susan Carey, Steve Pinker, Manizeh Khan, Mahesh Srinivasan, Nathan Winkler-Rhoades, and Rebecca Nappa 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).

ABSTRACT

While the referent of a non-reflexive pronoun clearly depends on context, the nature of these contextual restrictions is controversial. The present study seeks to characterize 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…)*. These differences have been attributed to: arbitrary lexical features encoding the bias of each verb, inferences based on nonlinguistic knowledge about the typical causes of different events, and systematic differences in the semantic structures of verbs. These experiments explore the semantic-structure hypothesis. In Experiment 1, we measured implicit causality judgments for the 720 highest-frequency transitive verbs in English. These verbs were grouped into classes based on an independently-motivated theory of verb meaning. Implicit causality judgments varied systematically with verb class. In Experiment 2, we focused on two verb classes encoding psychological states and tested all 264 verbs in these classes. Implicit causality judgments for the two classes 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.

Key words: pronoun resolution, implicit causality, thematic roles, psychological predicates, psych verbs

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).

(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 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, (3) and (4) 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. Brown and Fish (1983a) found 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.

They argued that it is this change in causal perceptions across verbs that accounts for the shift in pronoun interpretation. For this reason, both 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 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 conceptualization of social relationships and dominance hierarchies (Corrigan, 2001; LaFrance, Brownell, & Hahn, 1997; Maass, Salvi, Arcuri, & Gun 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 (Garnham, Traxler, Oakhill, & Gernsbacher, 1996; Greene & Gail McKoon,

1995; Guerry, Gimenes, Caplan, & Rigalleau, 2006; Koornneef & Van Berkum, 2006; Long & De Ley, 2000; McDonald & MacWhinney, 1995; McKoon, Greene, & Ratcliff, 1993; Shen & Yang, 2006; Stewart, Pickering, & Sanford, 2000; Pyykkonen & Jarvikivi, 2010) 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 (Frank, Koppen, Noordman, & Vonk, 2007; Kehler, Kertz, Rohde, & Elman, 2008; Crinean & Garnham, 2006; Ehrlich, 1980; Pickering & Majid, 2007; Stewart, Pickering, & Sanford, 1998). In the context of the Sapir-Whorf hypothesis (Whorf, 1956), researchers have asked whether IC is an effect of language on thought or of thought on language (Brown & Fish, 1983a; Hoffman & Tchir, 1990). 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. Garvey, Caramazza and colleagues (Garvey et al., 1974; Garvey & Caramazza, 1974; Caramazza, Grober, Garvey, & Yates, 1977) suggested that the IC bias is a function of the underlying semantics of the verb. Brown and Fish (1983a) extended this account by systematically relating each verb's bias (subject-biased, like *frighten*, or object-biased, like

fear) to the semantic roles of the verb's arguments. They distinguished four semantic roles and a total of three types of verbs: action verbs, stimulus-subject verbs, and stimulus-object verbs..

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, pronouns 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*). For both types of psych verbs, Brown and Fish argued that the pronoun is biased towards the STIMULUS. Thus the *stimulus-subject* verbs (*frighten, confuse, surprise*) are typically subject-biased, whereas *stimulus-object* verbs (*like, love, hate*) are typically object-biased – presumably because STIMULI are more causal than EXPERIENCERS.

Brown and Fish's account was attractive because, unlike previous accounts, it made predictions about which verbs would elicit which biases. However, subsequent research quickly discovered numerous cases in which these predictions were wrong. For example, *thank* and *punish* are predicted to be subject-biased but are in fact object-biased (Au, 1986). Several modifications to Brown and Fish's verb classification scheme were proposed, but all made similar errors (see Rudolph & Forsterling, 1997, for a review).

---

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

Given these repeated failures to link the IC pronoun bias to independently-motivated verb classes, this *semantic structure* approach was largely abandoned by the mid-1990s (but see Crinean & Garnham, 2006). Many subsequent researchers were 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; Maass, Salvi, Arcuri, & Gun Semin, 1989; Mannetti & De Grada, 1991; LaFrance, Brownell, & Hahn, 1997). These observations led many researchers to propose that IC is a result of non-linguistic knowledge about who tends to cause certain types of events (e.g., Au, 1986; Koornneef & Van Berkum, 2006; Maass et al., 1989). On this *world knowledge* account, IC biases themselves are not linguistic though they can influence language comprehension.

The *semantic structure* account consisted of two related claims: 1) the IC bias is linked to the verb itself (rather than being an emergent property of the utterance), and 2) the IC bias of a verb is predictable in some non-circular way from its meaning. The *world knowledge* account abandons the first claim, arguing that the IC bias of a sentence is the result of rapid calculations of the casual structure of events in the world.

Another alternative to the semantic structure hypothesis is the *arbitrary semantic tag* account, which keeps the first claim but abandons the second. On this hypothesis, verbs like *fear* and *frighten* do systematically give rise to distinct pronoun resolution biases, but the direction of the bias is not predictable based on independent semantic 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 appears to be

close to the original proposal in Garvey and Caramazza, 1974. See Hoffman and Tchir, 1990, for a related possibility).

Revisiting semantic structure

Despite its waning popularity, there are reasons to revisit the claim that semantic structure plays a role in pronoun resolution. First, there is little if any evidence *against* semantic structure playing a role in IC biases. The fact that non-linguistic factors such as social dominance affect IC biases has been taken as evidence in favor of the *world knowledge* account, but data of this kind are not inconsistent with the possibility that semantic structure plays an independent role in generating causal inferences.

The most serious problem for the *semantic structure* account is that the previously proposed versions of this hypothesis all made incorrect predictions. However, these early proposals invoked quite simple semantic structures resulting in only 3-4 semantic verb classes. The twenty-seven years since Brown and Fish (1983a) have witnessed an explosion of research in semantics, with many semanticists arguing for considerably more complex and nuanced accounts of the semantic structure of verbs (Goldberg, 1995; Jackendoff, 1992; Levin, 1993; Pinker, 1989; Pustejovsky, 1995; Talmy, 2000a, 2000b; Van Valin, 2004; see Levin & Rappaport Hovav, 2005, for a review). These new theories have been more successful at explaining a range of semantic phenomena (e.g., argument realization) than were earlier coarse-grained schemata like the one employed by Brown and Fish (1983a). It stands to reason that they may also better predict IC biases.

The present study

The present paper re-examines the semantic structure hypothesis by looking at the relationship between the semantic structures of verbs and their IC biases. To date our

knowledge of IC has been based largely on a limited number of verbs. Rudolph and Forsterling (1997) conducted a meta-analysis of the IC literature. Although dozens of studies had been conducted and 4 different languages had been tested, they were only able to include 256 individual verbs in the analysis, because many of the studies employed the same small set of verbs. This narrow focus increases the risk of over-fitting the data. Thus, in Experiment 1, we elicit IC pronoun resolution biases for 720 high-frequency verbs, nearly three times the number analyzed by Rudolph and Forsterling (1997). We then classify the verbs into semantic subclasses on the basis of independently-motivated theory of verb argument structure (Kipper, Korhonen, Ryant, & Palmer, 2008; Levin, 1993) and find that several large verb classes show systematic IC biases.

In Experiment 2, we test all of the verbs in two of these independently-motivated verb classes and again find essentially non-overlapping distributions, further suggesting that at least in some cases, the IC bias is directly related to verb class.

Experiment 1

In Experiment 1, we elicited the IC pronoun-resolution biases of 720 high-frequency verbs that allow animate subjects and objects. In contrast with previous studies, we did not pre-select the verbs because we believed that they had strong IC biases (Rudolph & Forsterling, 1997). Thus these data provide an unbiased estimate of the distribution of IC pronoun biases across verbs. This allows us to explore two questions. First, we can determine whether the distribution of IC biases is bimodal, consisting only of subject-biased and object-biased verbs, as seems to be assumed by at least some authors (e.g., Brown & Fish, 1983a), or continuously distributed, with many non-biased verbs (Semin & Fiedler, 1991; Garvey et al., 1974). Second, we divide the verbs into independently-

motivated semantic classes and examine whether verbs in the same class show the same pattern of IC bias. On the *semantic structure* account, it is expected that IC biases will systematically vary with verb class. This systematicity is not predicted by either the *world knowledge* or *arbitrary semantic tag* accounts, since neither posit a necessary connection between IC bias and the linguistic structure of the verb. Thus, on these theories, ancillary hypotheses would be required to explain any systematic relationship between semantic class and IC bias (see General Discussion).

Method

Participants: Participants were recruited and tested online through coglanglab.org. We analyzed the data from the 1,365 participants who were native English speakers, reported not 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 representative and theoretically unbiased, 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). We chose the 720 most common English verbs that met these criteria (Frances and Kucera, 1982).

Each subject was tested on 25 verbs randomly sampled from the total set. In order to minimize 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*. On each trial, the participant indicated by a mouse click which character was the dax. Participants were told a dax is a type of person,

but given no more information. 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 recognize 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)*.

Results and Discussion

Each verb was evaluated by an average of 47 participants. The distribution of results is shown in Figure 1. Results for each verb are presented in Appendix A. There was a slight overall bias in favor of the object (overall object bias: 59.2%, SE=0.6%, $t(719)=15.93, p<.01$). 43 verbs exhibited a significant subject bias (p's≤.05) and 272 a significant object bias (p's≤.05). Of these, 4 subject-biased and 84 object-biased verbs survive a conservative Bonferroni correction for 720 comparisons.

First we explored the distribution of IC biases. If IC bias is a binary distinction, with all verbs being either subject-biased or object-biased, then we would expect a bimodal distribution with most verbs clustering around the endpoints. Instead we found that the distribution of biases appears to be unimodal, suggesting the IC bias is continuous property and that many verbs have no strong bias. This was confirmed using the Hartigan dip test (.009, $p>.9$; Hartigan & Hartigan, 1985) implemented in R (Maechler, 2009; R Development Core Team, 2009). This is consistent with previous reports that many verbs do not elicit a systematic pronoun bias (Garvey et al., 1974; Semin & Fiedler, 1991).

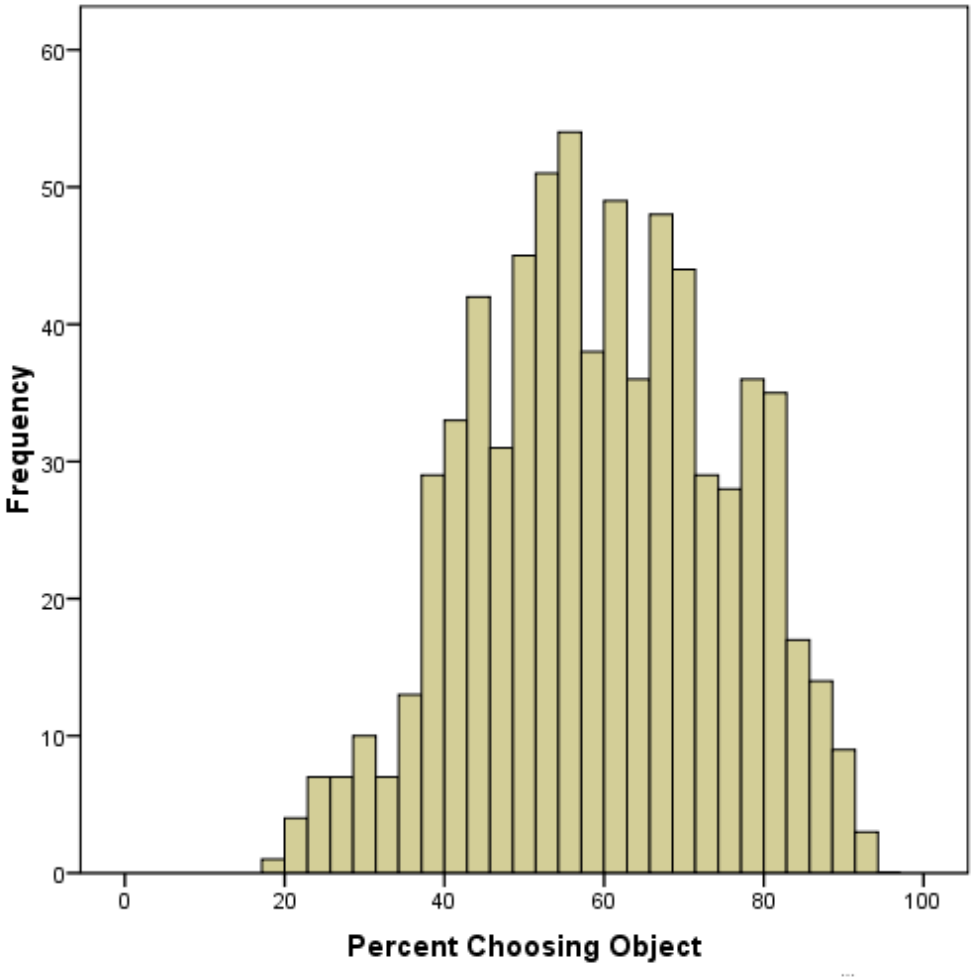

Figure 1. Histogram of IC biases elicited for the 720 verbs in Experiment 1.

VerbNet Analyses

On the hypothesis that IC is a function of verb meaning, we turned to an independent system of classifying verbs based on their semantic structure: VerbNet (Kipper et al., 2008). VerbNet is based on Levin's (1993) analysis of verb argument structure, supplemented with refinements suggested by Korhonen and colleagues (Korhonen & Briscoe, 2004; Korhonen & Ryant, 2005). Many proposals have been made about how to analyze verb meanings (Goldberg, 1995; Jackendoff, 1992; Levin, 1993;

Pinker, 1989; Pustejovsky, 1995; Talmy, 2000a, 2000b; Van Valin, 2004). While these proposals vary in the semantic distinctions that they address and the theoretical apparatus that they employ, many of the same classes of verbs appear repeatedly across different theories (see Levin, 1993).

Apart from any theoretical considerations, VerbNet offers notable advantages over other approaches. VerbNet is the most comprehensive semantic classification scheme for verbs, classifying 3769 verbs (including our 720) into 274 classes, and is available in electronic format on the Internet (verbs.colorado.edu/~mpalmer/projects/verbnet.html). Moreover, the classifications are built largely on objective, observable data (following Levin, 1993). First, verbs are distinguished based on the types of arguments they can take. Thus *tell* and *talk* are in classes 37.2 and 37.5, respectively, because the former can take a direct object (*tell him*) whereas the latter cannot (*talk him*). Second, verbs are distinguished based on what alternation classes they belong to. For instance, some verbs can appear both transitively and intransitively (*John broke the vase. The vase broke*), whereas others cannot (*John fell*). Finally, verbs were distinguished based on what the authors 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.[2]

(10) The fairy turned the straw to gold.

(11) The missionaries converted the natives to Deism.

---

[2] While this third criterion relies on intuitions about meaning, these judgments appear to be robust and this criterion is only employed in a small minority of cases.

In classifying the 720 verbs, we noticed that many of them are polysemous. For instance, *Mary touches Sally* can be interpreted as a contact event or as a psychological event. If the IC bias is the function of the verb's meaning, then it should depend on which meaning was assigned to the verb by the participants. If different participants select different interpretations of polysemous verbs, and the different interpretations have contrasting biases, IC effects could be masked. Interestingly, despite the ubiquity of verbal polysemy, this issue does not appear to have been addressed previously in the IC literature.

To address this, for each verb we tallied all syntactically-appropriate verb classes given by VerbNet and then excluded all verbs with more than one possible use. Specifically, we limited our analysis to verb classes that appear in the transitive frame. Thus, the intransitive use of *break* (*The vase broke*) or the preposition-requiring sense of *worry* (*Sally worried* about *Mary*) did not figure into this analysis. We included classes not licensed for two animate arguments.[3]

---

[3] There was one exception to this policy: VerbNet further divides the 274 verb classes into 5,257 senses. Most verb class subdivisions reflect syntactic as well as semantic distinctions (e.g., *Mary sank the ship* vs. *The ship sank*). However, some subdivisions reflect different semantic roles for the arguments. For example, in *John dried the clothes*, the subject is an AGENT, whereas in *The hairdryer dried the clothes*, the subject is an INSTRUMENT. However, since these distinctions seem to focus on the animacy of the grammatical subject (e.g., John), and all of our subjects are animate, we did not treat these cases as polysemous (N=49). In any case, fully 29% of these verbs showed a significant bias one way or another, so it was not the case that these verbs constituted the bulk of non-biased verbs.

After excluding all verbs with more than one possible use, the remaining 317 monosemic verbs again showed a broad distribution of biases with a slight overall object bias (58.9% choosing object, SE=0.9%, $t(316)=5376.16$, $p<.01$). The distribution of results is shown in Figure 2. Again, there is no evidence of a bimodal distribution (Hartigan's *dip*=0.015, $p>.6$). Thus, even considering verbs with only one possible meaning consistent with the syntax, 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 A).

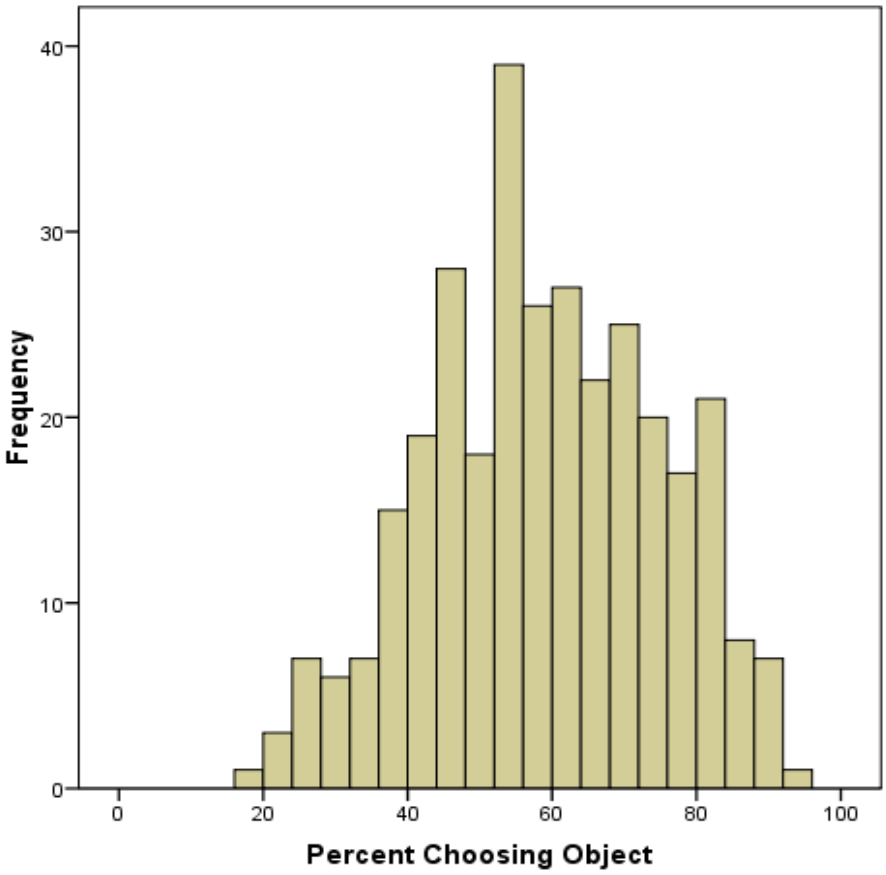

Figure 2. Histogram of IC biases elicited by the 317 monosemic verbs.

We then asked whether IC biases vary systematically with the verb classes identified by VerbNet system. We found four verb classes for which significant numbers of monosemic verbs (N>10) had been tested (see Table 1). Since our full set of 720 verbs showed a slight overall bias in favor of the object (59.2%), the most conservative test is not whether the verbs in a given class differ significantly from a 50% chance level, but whether membership in a given verb class predicts IC bias significantly better than the grand mean (59.2%). Thus, we conducted statistical tests in terms of difference from grand mean.

By this criterion, all four verb classes were significantly predictive of IC bias (class 45.4 is not significantly different from 50%, but the remaining three are.) Thus we find that IC pronoun bias varies systematically across independently-defined semantic verb classes. Such results are predicted by the *semantic structure* account but not clearly predicted by the *arbitrary semantic tag* account or the *world knowledge* account.

Class	Verbs	Object Bias (SE)	Significance
45.4	accelerates, balances, cools, corrects, doubles, dries, enhances, enlarges, improves, intensifies, obscures, reproduces, reverses, revives, shortens, shrinks, shuts, sinks, slows, softens, strengthens, wakens, weakens, widens	51.1% (2.2%)	$t(23)=3.56$, $p<.01$
31.1	affects, arouses, bores, calms, confuses, discourages, disturbs, entertains, frustrates, offends, pleases, puzzles, satisfies, surprises, threatens, troubles, worries	36.0% (3.3%)	$t(16)=6.39$, $p<.01$
31.2	admires, appreciates, bears, cherishes, despises, dislikes, favors, fears, hates, loves, misses, resents, respects, suffers, worships	75.3% (4.2%)	$t(14)=3.88$, $p<.01$
33	attacks, blames, blesses, celebrates, commends, condemns, congratulates,	72.5% (3.0%)	$t(14)=4.38$, $p<.01$

	criticizes, curses, denounces, excuses, forgives, greets, honors, thanks

Table 1. Implicit causality biases for the four verb classes with at least 10 tested verbs.

Experiment 2

Experiment 1 provided preliminary evidence that the IC bias varies systematically across fine-grained semantic verb classes. In Experiment 2, we explored whether this pattern generalizes to all verbs of a given semantic class. We focused on class 31.1 (*frighten, confuse, surprise*) and 31.2 (*fear, love, respect*). These classes contain the bulk of transitive psych verbs, which have been a focus of attention in the IC literature since Brown and Fish (1983a) noticed that many STIMULUS-EXPERIENCER verbs (in which the subject evokes a mental state in the object: e.g., class 31.1) are subject-biased, while many EXPERIENCER-STIMULUS verbs (in which the object evokes a mental state in the subject: e.g., class 31.2) are object-biased.

Thus, in Experiment 2, we tested participants on all members of classes 31.1 and 31.2 from Levin (1993), using the same method as in Experiment 1. On the *semantic structure* account, we expect the IC bias to vary systematically with verb class. Such results would not be predicted by either the *arbitrary semantic tag* account and the *world knowledge* account.

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: The materials and procedure were identical to Experiment 1, except that each participant was presented with 12 verbs randomly selected without replacement from each of two verb classes: 33.1 ("frighten" verbs; N=220) and 33.2 ("fear" verbs; N=44).

Results and Discussion

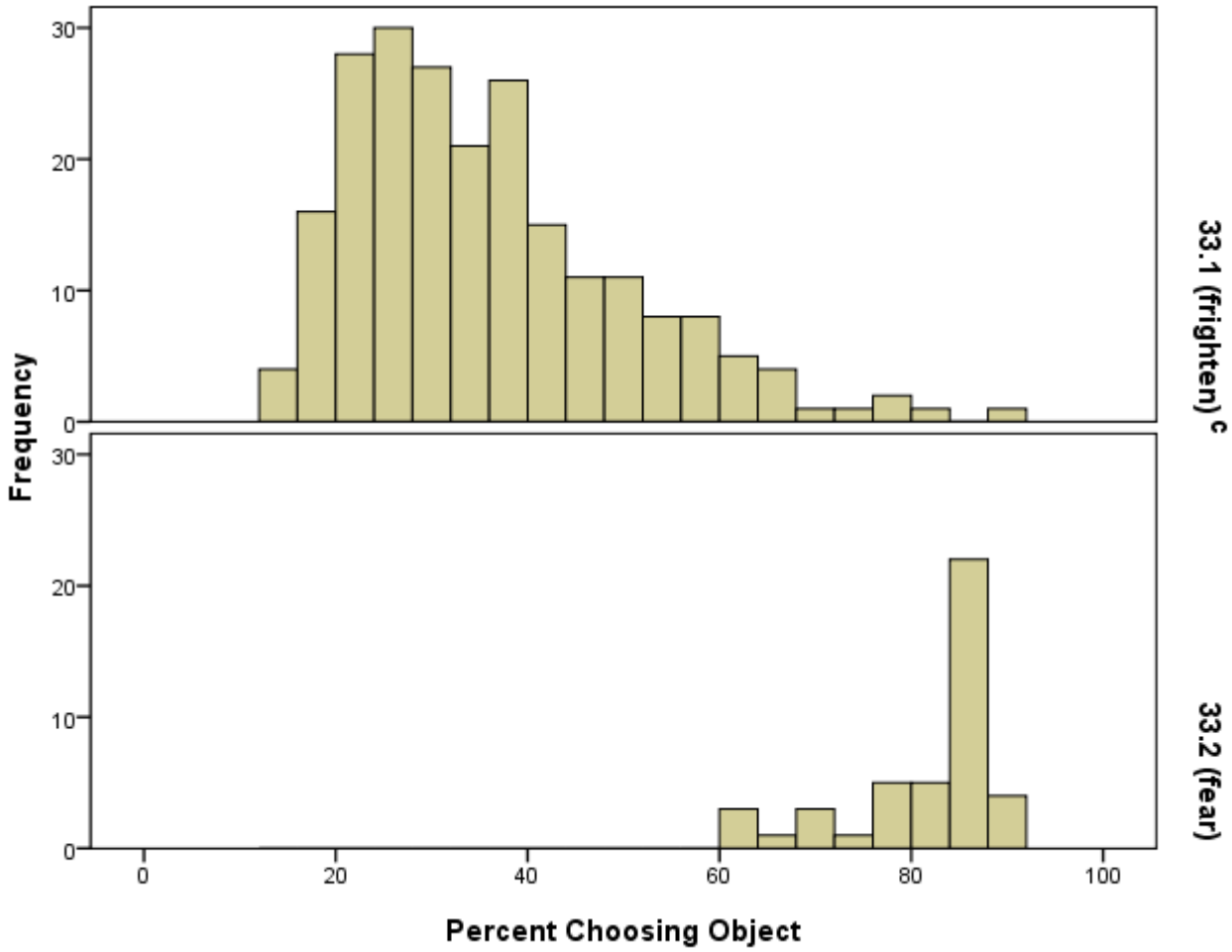

Figure 3. Histogram of IC biases for all "frighten" (class 33.1) and "fear" (class 33.2) verbs tested.

Individual verb results are presented in Appendix B; the distributions are shown in Figure 3. In contrast to Experiment 1, the distribution was clearly bimodal (Hartigan's *dip*=0.036, $p<.05$). *Frighten* verbs (33.1) showed a strong subject bias (35.7% object bias, SE=1.0%, $t(219)=24.44$, $p<.01$), while *fear* verbs (33.2) showed a strong object bias (81.5% object bias, SE=1.1%, $t(43)=20.18$, $p<.01$), and the two classes were significantly different from one another $t(262)=20.72$, $p<.01$.

Again, using the grand mean of 59.2% from Experiment 1 as our conservative "chance" threshold, all *fear* verbs (33.2) exhibited object-biases, 41/44 significantly so (39/44 after Bonferroni correction), while 204/220 *frighten* verbs (33.1) exhibited subject-biases, 171 significantly so (112 survive Bonferroni correction).[4] Only 6 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*=0.042, $p<.02$). Of the remaining 161 *frighten* verbs, all but 4 (*wounds, dejects, cows, alienates*) exhibited a subject bias, 142 significantly (83 after Bonferroni correction). Only one *frighten* verb (*alienate*) was significantly object-biased. All of the 34 remaining *fear* verbs

[4] All 44 *fear* 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 *frighten* verbs still show a subject bias (130 significantly), 34 an object bias (12 significantly), and 2 no bias.

were significantly object-biased (33/34 after Bonferroni correction). Moreover, there was no overlap between the distributions of the two classes, with the exception of *alienate*.

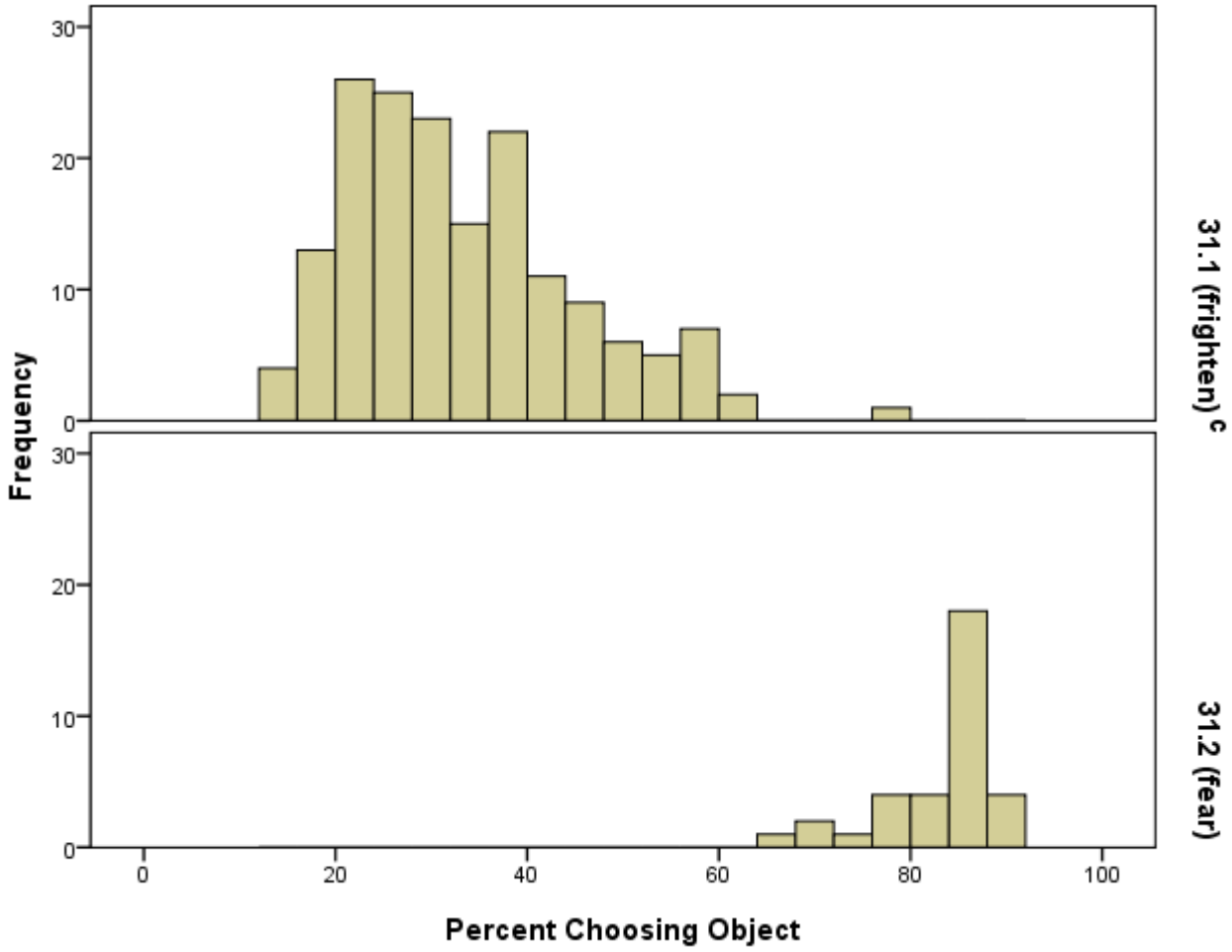

Figure 5. Histogram of IC biases for monosemic "frighten" (class 33.1) and "fear" (class 33.2) verbs.

General Discussion

Experiments 1 and 2 demonstrate that implicit causality (IC) biases in pronoun interpretation vary systematically across semantic classes of verbs that are independently motivated based on patterns of argument realization. In Experiment 1, we investigated four different verb classes, finding significant biases for each. In Experiment 2, we investigated two of these classes (*fear* and *frighten* verbs), finding that the IC bias was consistent for the

vast majority of members in both classes. These data revive the *semantic structure* hypothesis, which directly predicts a systematic relationship between semantic structure and IC bias. These results are not directly predicted or explained by alternate accounts.

In the remainder of this discussion, we explore three questions: First, we examine earlier *semantic structure* accounts of IC and show that they fail to make the correct distinctions and predictions. Second, we explore how the semantic structures that underlie the verb classes in VerbNet could potentially support causal inferences. Finally, we revisit the *world knowledge* and *arbitrary semantic tag* accounts to explore how they might account for these findings.

Semantic roles vs. semantic structures

Previous *semantic structure* accounts (e.g., Brown & Fish, 1983a) were based on the linguistic concept of a thematic (or semantic) role. Thematic roles are discrete categories encoding the role that an entity can play in the event denoted by a verb. Brown and Fish distinguish four thematic roles: AGENT, PATIENT, STIMULUS, and EXPERIENCER. On this account, psych verbs involve a STIMULUS and an EXPERIENCER. For some verbs, the STIMULUS is the subject (16) whereas for others, the EXPERIENCER is the subject (17). All other transitive verbs involve an AGENT and a PATIENT; the AGENT is always the subject (18-19).

(16) Mary$_{\text{STIMULUS}}$ frightened Sally$_{\text{EXPERIENCER}}$.

(17) Bernard$_{\text{EXPERIENCER}}$ admired Alfred$_{\text{STIMULUS}}$.

(18) Susan$_{\text{AGENT}}$ ate the sandwich$_{\text{PATIENT}}$.

(19) John$_{\text{AGENT}}$ broke the vase$_{\text{PATIENT}}$.

Brown and Fish argued that the roles STIMULUS and AGENT are intrinsically more causal than the roles THEME and EXPERIENCER, leading to systematic patterns of IC bias.

However, as we saw earlier, this account made predictions about IC bias which were found to be erroneous. Our data set confirms that this is the case. Numerous verbs classified as AGENT-PATIENT by Brown and Fish are in fact object-biased, including class 33 as discussed above (see Table 2). Brown and Fish also make incorrect predictions about some psych verbs, such as *comprehend*, *understand*, *discern*, *sense*, *hear* and *detect*, which are EXPERIENCER-STIMULUS verbs and thus predicted to be object-biased, but in fact are subject-biased (see Appendix A). Note that VerbNet does not include these verbs in class 31.2; rather, these verbs belong to one of two smaller classes of psych verbs (87.2 and 30.1).

Brown and Fish used a very small number of thematic roles to capture the relevant aspects of verb meaning. One might wonder whether a more fine-grained thematic role theory could correctly predict the pattern of IC biases, making the use of semantic verb classes (defined by argument realization patterns) unnecessary. In addition to the semantic verb classes, VerbNet categorizes the arguments of each verb into 21 discrete thematic roles. We reanalyzed the data from Experiment 1 with respect to these roles. The monosemic verbs from Experiment 1 represented 21 different combinations of thematic roles (with a total of 5 different roles in subject position and 16 in object position)[5] Only six of these combinations were represented by more than ten verbs (see Table 2). Once again two of these thematic role combinations (CAUSE-EXPERIENCER and EXPERIENCER-THEME) picked out the large classes of psych-verbs (class 31.1 and 31.2 respectively), which have

[5] The 49 verbs in 8 verb classes which allow more than one syntactically-appropriate argument structure (see footnote 3) were not included for these analyses. Also excluded were 30 verbs for which VerbNet did not provide a full semantic role analysis.

systematic IC biases. The mean biases for the other four combinations, however, were not reliably different from the grand mean (59.2%), even though these classes contained many groups of verbs with strong IC-biases. For example, the AGENT-PATIENT set included the subject-biased *dry* verbs (class 45.4), while the AGENT-THEME set included the object-biased judgment verbs (class 33). Thus thematic role combinations only predict IC-biases when they pick out the same verb classes as finer grained theories of semantic structure (e.g., classes 31.1 and 31.2). Where thematic roles provide a coarser categorization of verbs they fail to capture distinctions in meaning which influence IC pronoun bias (e.g., classes 45.4 and 33).

Semantic Roles	# of Verbs	Object Bias (SE)	Significance
AGENT-THEME	138	61.6% (1.3%)	$t(129)=1.78$, $p=.09$
AGENT-PATIENT	34	57.4% (2.2%)	$t(32)<1$
CAUSE-EXPERIENCER	17	36.0% (3.3%)	$t(16)=6.39$, $p<.01$
EXPERIENCER-THEME	15	75.3% (4.2%)	$t(14)=3.88$, $p<.01$
AGENT-TOPIC	14	62.3% (3.0%)	$t(10)<1$
AGENT-RECIPIENT	12	58.8% (3.2%)	$t(11)<1$

Table 2. Implicit causality biases for the six argument structures with at least 10 tested verbs.

Thus, even a more comprehensive thematic role theory is insufficiently fine-grained to make the right predictions. Interestingly, thematic role theories also failed to solve the problem for which they were designed: explaining why particular arguments surface in particular syntactic positions (for a comprehensive review of the topic, see Levin & Rappaport Hovav, 2005). For this reason, many theorists have abandoned thematic role hierarchies and instead conceive of argument realization as a process of mapping between a structured semantic representation and syntax. On such a theory, thematic roles are merely an emergent phenomenon defined by the position of an argument in the semantic

structure of the verb. The verb classes in VerbNet instantiate this kind of theory; they are defined by patterns of argument realization and each one has an associated semantic structure. By analyzing our data in relation to these verb classes we implicitly adopted the notion that the semantic structure of a verb is the relevant unit for determining IC bias, and not its set of thematic roles.

Semantic structures and implicit causality

On the thematic role account, the IC bias of verb was thought to reflect the degree of causal agency attributed to each role. If IC bias is a function of semantic structures, rather than thematic roles, what property of these structures shapes the causal bias? What is it about the four semantic classes identified in Experiment 1 that leads the verbs in them to be subject- or object-biased? For each verb class, VerbNet provides a semantic structure which decomposes the verb into more primitive predicates that assign functions to the arguments of the verb. These semantic structures are directly modeled on Moens and Steedman's theory of event decomposition (1988) but they share many features with a wide range of proposals in lexical semantics (e.g., Harley, 1995; Jackendoff, 1992; Pinker, 1989).

These semantic structures suggest some hypotheses about the semantic distinctions that underlie IC biases. For both of the subject-biased classes—the *frighten* verbs (31.1) and the *dry* verbs (45.4)—VerbNet provides a structure in which the subject of the sentence is identified as the cause of a sub-event. The object-biased *fear* (31.2) class is amenable to a similar explanation. While some theorists have suggested that the object of *fear* verbs is also a cause (Jackendoff, 1990; Pinker, 1989), VerbNet does not analyze 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.[6] Levin (1993) suggests a similar structure for class 33 (*blame*, *admire*, etc., see also Greene and McKoon, 1995):

> These verbs share some properties with the *admire*-type psych-verbs [e.g., *fear* 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. (p. 196)

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*. Thus, in all four of these cases, the semantic structure provides clear information about the entity that is causally responsible for the event. This information, in combination with expectations about the content of subordinate clauses 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.[7]

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

---

[6] " EMOTIONAL_STATE(E, EMOTION, EXPERIENCER) IN_REACTION_TO(E,THEME)"

[7] Early discussions of of IC implicitly assumed that a subordinate clause introduced by *because* necessarily encodes the cause of the event in the main clause. More recently, Pickering and Majid (2007) have suggested that *because* introduces an explanation, rather than a cause. Either account is consistent with the analysis here, since explanations by necessity are more likely to refer to the causal than non-causal entity.

differences in the conceptualization 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, this does not mean that these theoretical distinctions are vacuous or not empirically testable. 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 and need not be calculated on the basis of world knowledge. Because IC is predicted by another linguistic property (semantic structure), the learner can infer the IC of a verb without hearing sentences in which the cause is explicitly labeled. In contrast, on the *arbitrary semantic tag* hypothesis, learners acquire a feature for each individual verb that indicates which argument is typically assumed to be the cause of events described by that verb. This proposal also attributes IC to linguistic representations, but it would have to be acquired from utterances in which the cause is explicitly marked (e.g., “Mary daxed Alfred because he was a jerk”). Finally, the

world knowledge account argues that IC effects do not depend directly on linguistic representations but instead reflect broad inferences that we make about the events that are described in a sentence on the basis of our prior knowledge of similar situations. Linguistic experience might not be necessary, but experience with the world is: stable IC biases would depend on the degree to which we had encountered similar events and had been able to infer their causes. 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 knowledge of semantic structure, linguistic experience and event knowledge. However, no such studies have been conducted.

With all this in mind, it is worth revisiting the most commonly-cited evidence in favor 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). These studies involved offline judgments of causality employing the Brown and Fish (1983a) paradigm. In a typical study, participants read a sentence like (20) and determine whether this event occurred because Ted is the sort of person who criticizes people or because Paul is the sort of person whom people criticize (in this case, people would rate the latter possibility more highly since *criticize* is an object-biased verb).

(20) Ted criticizes Paul.

The crucial manipulation for these later studies was to vary the social relation between the two characters. For instance, Corrigan (2001) found that participants were more likely to declare the object responsible in sentences like (21) than in sentences like (22), presumably because traitors are seen as more deserving of criticism than kings.

(21) The monarch criticized the traitor.

(22) The traitor criticized the monarch.

How is information about criticism events integrated with information about monarchs and traitors? On the *world knowledge* account, listeners access statistical knowledge about why monarchs and traitors criticize one another to ultimately decide who was the most likely cause. On the *semantic structure* account, listeners first build a semantic structure based on meaning of the verb. For a verb like *criticize*, the grammatical object (*traitor* in 21, *monarch* in 22) is semantically marked as the cause. 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, linguistically-derived assessment. A similar process would be expected under the *arbitrary semantic tag* account: an initial linguistic analysis of the grammatical object as the cause, which is later modified by world knowledge.

Since it is possible to track moment-to-moment adjustments in listener's online assessments of pronoun resolution (Arnold, Eisenband, Brown-Schmidt, & Trueswell, 2000; Arnold, Brown-Schmidt, & Trueswell, 2007; 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. At the present time we do not know whether the social variables which influence explicit judgments actually have an effect on pronoun resolution, let alone when this effect emerges during language comprehension. Since the *world knowledge* account attributes both verb-specific factors and effects of social variables to the same underlying process, it necessarily

predicts that the social factors would influence pronoun resolution. The other theories need not make such a claim.

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 3 or 4 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, 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 linguistic structure, they may provided 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 from 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 more than quadruples the number of English verbs for which IC biases have been reported (*cf* Rudolph & Forsterling, 1997). By making these data publically available (see the appendixes) we hope to facilitate the creation of new experiments (see Goikoetxea, Pascual, & Acha, 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.

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.

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.

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.

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

semantic roles. *Language and Cognitive Processes*, *21*, 636-648.

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.

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

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.

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

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.

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.

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.

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. In *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 University Press: Cambridge, UK.

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.

Maass, A., Salvi, D., Arcuri, L., & Semin, G. (1989). Language use in intergroup contexts: The

linguistic intergroup bias. *Journal of Personality and Social Psychology*, *57*, 981-993.

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.

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

Pinker, S. (2007). *The stuff of thought: Language a window into human nature*. New York: Viking.

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

Rudolph, U., & Forsterling, F. (1997). The Psychological Causality Implicit in Verbs: A Review. *Psychological Bulletin*, *121*, 192-218.

Semin, G. R., & Fiedler, K. (1991). The linguistic category model, its bases, applications and range. In *European review of social psychology* (pp. 1-30). Chichester, England: Wiley.

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

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.

Stewart, A. J., Pickering, M. J., & Sanford, A. J. (1998). The influence of verb bias information on clausal integration: Implicit causality and implicit consequentiality. Presented at the Eighth Annual Meeting of the Society for Text and Discourse, Wisconsin.

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. (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 A: 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 B: 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%
astonishs	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%