Greatest surprise reduction semantics: an information theoretic solution to misrepresentation and disjunction

D. E. Weissglass

Abstract Causal theories of content, a popular family of approaches to defining the content of mental states, commonly run afoot of two related and serious problems that prevent them from providing an adequate theory of mental content—the misrepresentation problem and the disjunction problem. In this paper, I present a causal theory of content, built on information theoretic tools, that solves these problems and provides a viable model of mental content. This is the greatest surprise reduction theory of content, which identifies the content of a signal as the event the surprisal of which is most reduced by that signal. Conceptually, this amounts to the claim that the content of a signal is the event the probability of which has increased by the largest proportion, or the event that the signal makes the most less surprising to us. I develop the greatest surprise reduction theory of content in four stages. First, I introduce the general project of causal theories of content, and the challenges presented to this project by the misrepresentation and disjunction problems. Next, I review two recent and prominent causal theories of content and demonstrate the serious challenges faced by these approaches, both clarifying the need for a solution to the misrepresentation and disjunction problems and providing a conceptual background for the greatest surprise reduction theory. Then, I develop the greatest surprise reduction theory of content, demonstrate its ability to resolve the misrepresentation and disjunction problems, and explore some additional applications it may have. Finally, I conclude with a discussion of a particularly difficult challenge that remains to be addressed—the partition problem—and sketch a path to a potential solution.
1 Misrepresentation and the disjunction problem

The primary project of psychosemantics is to identify those properties that fix the content of mental states. This project has a long history of varied approaches, but contemporary work has largely settled on a causal approach to mental content—one which aims to identify the content of a mental state in terms of its actual, typical, or proper cause.¹ Causal theories of content share the intuition that, though other factors may play a role, the apparent causal connection between states of the world and mental states is the most important consideration in understanding the content of mental states. Whatever else we may need to consider, the fact that a mental state ‘cat’ is tokened when a cat is present is an essential part of the story of content.

Causal theories of content have an intuitive appeal, but are persistently undermined by two problems that have been extraordinarily resistant to solution. The first of these is the misrepresentation problem. Causal theories have trouble making sense of cases where a mental state is tokened despite the absence of its (proper) cause, and so have trouble accounting for instances where our mental states misrepresent the world. The second is the disjunction problem. Causal theories are prone to assigning a disjunction of all possible causes of a mental state to its content.

The simplest demonstration of these problems is given by Fodor (1987), who presents a toy theory called the ‘crude causal theory’ (CCT). The CCT states that the content of a mental state is the set of properties that reliably causes the state to be tokened:

\[\text{(CCT): The content of a representation R for subject S is the property or set of properties P such that P reliably causes R in S}\]

That is, the content of the mental state ‘horse’ is a horse just because that mental state is reliably caused by the property (or properties) of horseness (Fodor 1987, pp. 99–102).

Like causal theories of content in general, the CCT is intuitive but troubled by both the misrepresentation and disjunction problem. The CCT cannot distinguish apt from inapt reliable causes of a mental state, and so cannot account for signals that misrepresent. ‘Horse’ tokens might be reliably caused by properties other than horseness—for instance, cowness, donkeyness, muleness, etc.—none of which are plausibly part of the content of the mental state ‘horse’. Likewise, CCT is vulnerable to the disjunction problem. If horseness and cowness are both reliable causes of ‘horse’, then the CCT seems to imply that the content of ‘horse’ is not horseness or cowness but the disjunction of the two (Fodor 1987, pp. 99–102).

The misrepresentation and disjunction problems are challenging for causal theories of content because they complicate its core intuition. It may be that the

¹ ‘Causal theories’ here includes teleosemantic approaches, hence the mention of ‘proper’ causes.
content of a mental state is somehow related to the causes of that state, but it cannot be
so simple as to say that the content of a mental state just is whatever causes it. Some
states may be caused on occasion by events totally unrelated to their causes. If we are to
have a successful causal theory of content, we will need to draw a line between the
causes that matter and the causes that don’t (Rupert 2008, pp. 356–357).

Despite wide acknowledgment of the seriousness of these problems, they remain
largely unresolved even in cutting edge of work in psychosemantics. Unresolved as
they are, the misrepresentation and disjunction problems threaten to undermine
causal theories of content, and perhaps the project of naturalistic psychosemantics at
large. In this paper, I develop the greatest surprise reduction model of content
(GSR), a causal theory of content that serves well as a general model of content
while being immune to both the misrepresentation and disjunction problems.
However, I will first review two contemporary causal theories of content and
demonstrate that both remain vulnerable to the misrepresentation and disjunction
problems, as well as other challenges. This serves not only to demonstrate the need
for a solution to these problems, as well as to establish some of the key concepts
needed to articulate the GSR.

2 Semantics, probability, and information

Information semantics is a particularly influential branch of causal theories of
content, which takes the content of a mental state to be the product of the
information that state carries about various events.2 The contemporary interest in
information semantics is largely due to Dretske (1981) and his work extending
information theoretic tools built to describe the average information content carried
by signals in a given channel of communication to describe the information carried
by particular signals. Dretske believed the result was a set of powerful tools for
understanding the amount of information carried by any individual signal, for
comparing how much information is carried by different signals, and various other
features important to the functioning of a signal (pp. 47–55).

The informational content of a signal, by Dretske’s account, exemplifies the basic
intuition of information semantics—that a signal carries the information that some event
has occurred is a product of the way that signal affects the probability of that event.
According to Dretske, a signal carries information about those events which had a pre-
signaling probability of less than one and post-signaling probability of one. For events the
probability of which is affected by a signal, but not raised to one, we can talk loosely of that
signal having the content that such an event occurs with the given probability, though
Dretske is skeptical of how useful this is (Dretske 1981, pp. 65–69).

Despite being responsible for the popularity of information semantics, Dretske’s
own view of what information theory could do is quite limited. Dretske was
convinced from the start that information theory will not be able to provide us with a
complete semantic theory, and restricts its application to a related but distinct notion

2 I use ‘events’ not ‘states’ because this will applied to acts in handling directive signals in section three.
of semantic information (1981, pp. 40–46). Likewise, the limits of Dretske’s own model in handling less than perfect information are troubling, as imperfect information is more the rule than the exception in natural systems.

However, recent approaches to information semantics are a bit more ambitious. Two related models, one from Skyrms (2010a, b) and the other from Shea, Godfrey-Smith, and Cao (2018), allow for the informational content of a signal to include events the post-signaling probability of which is less than one, and also take informational content to be capable of defining propositional content—the sort of meaning that Dretske thought information theory could never provide us. Reviewing these models will demonstrate the basic structure of information semantics, as well as the challenges it faces.

2.1 Skyrms/Birch information semantics

Like most information semantic theories, the basic building block of Skyrms’s theory of content is the information that a given signal carries about a given event in the world. Skyrms takes the information that a signal carries about an event to be determined by the ratio of the post-signaling probability of that event and the pre-signaling probability of that event. This is perhaps easiest to see in the formalism. Take \( I(s_1(e_1)) \) to denote the information that signal \( s_1 \) carries about event \( e_1 \), \( \Pr(e_1) \) is the probability of event \( e_1 \) within the signaling game, and \( \Pr(e_1|s_1) \) is the probability of event \( e_1 \) conditional on signal \( s_1 \). The information carried by \( s_1 \) about \( e_1 \) is then determined as follows:

\[
I(s_1(e_1)) = \log_2 \frac{\Pr(e_1|s_1)}{\Pr(e_1)}
\]  

Note that neither the pre-signaling nor post-signaling probabilities of an event will realistically reach zero, so we do not need to worry about the possibility of the information carried by a signal about an event being undefined (Skyrns 2010a, pp. 35–36).

Skyrms defines the total informational content of a signal as a vector of the information carried by that signal about each event about which the signaling system is able to communicate. The result is an \( n \)-dimensional vector, where each dimension of the vector represents a distinct event. Formally, the total informational content of a signal \( s_1 \) about all events \( e_1 \) to \( e_n \) in the set of events \( E \), denoted \( I_E(s_1) \) is calculated as follows (Skyrns 2010a, pp. 40–42):

\[
I_E(s_1) = \left< \log_2 \frac{\Pr(e_1|s_1)}{\Pr(e_1)}, \log_2 \frac{\Pr(e_2|s_1)}{\Pr(e_2)}, \ldots, \log_2 \frac{\Pr(e_n|s_1)}{\Pr(e_n)} \right>
\]  

Each component of this vector tells us how the signal changes the probability of one event. Taken together, we have a picture of how the signal changes the probability of each event in the signaling system.

---

3 Skyrms and Barrett (2019) provides a good discussion of the differences and similarities of these approaches, as well as Birch’s (2014) extension to Skyrms (Skyrns, 2010a, 2010b).
Consider a simple case of a signal sent which carries information about two events that each have a pre-signaling probability of 0.5. If signal $s_1$ raises the probability of event $e_1$ to 1, and reduces the probability of event $e_2$ to 0, the resulting informational content of that signal is:

$$I_E(s_1) = \begin{pmatrix} \log_2 \frac{1}{0.5}, \log_2 \frac{0}{0.5} \end{pmatrix} = \langle 1, -\infty \rangle$$

This vector says that the signal $s_1$ makes us absolutely certain that the first event, and not the second, obtains. Because of the properties of logarithms, the information carried about an event with a probability reduced to zero will always be a negative infinity.

Skyrms defines the propositional content of a signal in terms of its informational content. The propositional content of a signal is the set of events for which the signal carries more than $-\infty$ bits of information. This might sound a bit strange, but its translates to all those events that the signal does not rule out (i.e., those events with a non-zero post-signaling probability), which is similar to some fairly standard ways of thinking about propositions. In simple cases, like the example considered above, where a signal raises the probability of one event to one and reduces others to zero, this gives an intuitive result. The content of our simple example signal is $e_1$. In a more complex case, we might have any number of events with a non-zero post-signaling probability, which would all then qualify as part of the content of the signal (Skyrms 2010a, pp. 40–42).

Skyrms’s proposal demonstrates the intuitive appeal of information semantics, but also the challenges that have thus far prevented it from being successful. In its favor, Skyrms’s proposal takes a fairly lightweight set of mathematical tools, alongside some fairly uncontroversial observations about the probabilistic relations between signals and events, and produces a formally rigorous theory of content based on naturalistic properties. However, Skyrms’s model fails against the same challenges that have plagued nearly every other causal theories of content—it fails to answer the misrepresentation and disjunction problems. This failure, it turns out, produces some rather severe effects.

A problem for information semantics in general is that each signal carries information about a large number of events. Because a signal’s carrying information about some event is just a matter of that event having a non-zero probability conditional on that signal, the mere fact that some event occurs alongside a given signal is by itself a guarantee that the conditional probability of that event of the world given that signal is greater than zero. For Skyrms’s account, the fact that all events that occur alongside a signal have a non-zero post-signaling probability means that none of them are actually ruled out by the signal, and so all of them must be part of that signal’s propositional content. Not only does this have the strange implication that events made less likely by a signal are a part of its propositional content, it renders this approach extremely vulnerable to the misrepresentation and disjunction problems. For a signal to misrepresent requires exactly something that Skyrms’s account cannot allow—that it occurs alongside an event that is not part of its content. Likewise, the fact that all events which occur alongside a signal are part
of its content seems to result in enormously disjunctive content attributions (Birch 2014, pp. 499–500).

These problems are compounded if we follow Skyrms in noting that neither pre-signal nor post-signal probabilities in actual signaling systems are likely to actually hit zero. This means we should expect the propositional content of a signal to include all events in the signaling system, as none will be ruled out. We might try to repair this by using an error term that treats post-signaling probabilities near enough to zero as though they were zero, but this will likely not be sufficient, as it is unlikely that we will find a non-arbitrary error term that is sufficiently large to prevent rambling disjunctions. This is especially the case for signals that reliably or regularly misrepresent, as handling these might require an error term so large as to eliminate content altogether.

In light of these difficulties, Birch (2014) proposes a modification to Skyrms’s account, which is meant to solve the misrepresentation problem. Instead of taking the content of a signal to be a product of the information that signal carries in the current system, we can take it to be determined by its informational content at the nearest separating equilibrium—a counterfactual signaling system where all signals perfectly indicate the event observed by the signaler. The nearest separating equilibrium is an evolutionary concept, referring to a state of the signaling system at large when all signalers have adopted perfect one-to-one mappings of events to signals. As such, each signal at a separating equilibrium carries perfect information about the event observed by the sender, ruling out all but one event. The simple example considered in Eq. 3 is a case of such a separating equilibrium, and it may be because of this that Skyrms’s model gives such an intuitive solution to that example. Separating equilibria are evolutionarily stable, and so a system will remain in a separating equilibrium once it has arrived at it—unless acted on by significant outside pressure (Birch 2014, pp. 503–505).

Birch’s nearest separating equilibrium proposal purports to solve the misrepresentation problem by providing a way to specify contents even for signals that carry information about multiple facts in the current state of the convention by appealing to facts about how that system will or might evolve. The separating equilibrium that is nearest in terms of the evolutionary dynamics of the current signaling system is used to determine the content of a signal. That is to say that the content of a signal is a product of the information that signal would carry at the nearest iteration of the system in which all players of the game act in total accordance with their type—the signal sent tells us exactly what event a sender observed. This formally precise counterfactual notion is meant to allow us to say that a signal misrepresents when it carries information other than the information it would carry at the nearest separating equilibrium (Birch 2014, pp. 504–505)

Despite its sophistication, Birch’s proposal runs into a serious challenge in defining the critical notion of ‘nearness’. We need to know what ‘nearness’ means if we are to determine on which counterfactual signaling equilibrium we are to base our judgments of content. However, making ‘nearness’ itself into a precise notion turns out to be very difficult. As a result, it is not clear that we can determine which counterfactual equilibrium we should consider, and so the value of this approach is questionable.
Birch notes the difficulty in defining ‘nearness’ as it could mean any number of things—we might consider changes in population frequencies, the number of generations required to pass, etc. However, Birch hopes that the potential ambiguity about which equilibrium is nearest will not be too much of a problem. Hopefully, all of the equilibria that are ‘nearest’ by any plausible definitions will agree on the content of a signal, rendering this problem insignificant. When these cases do not agree, Birch suggests that we should be satisfied to say that there is no determinate nearest separating equilibrium—and therefore no determinate content (Birch 2014, p. 504).

However, it is not clear that Birch’s hope in the convergence of plausible nearnesses is justified, or even what all the plausible forms of nearness are. Without this having been demonstrated, we have little reason to believe that the nearest separating equilibrium model will be successful for the bulk of content attribution cases. This is particularly troubling as many actual conventions show no clear sign of evolution to separating equilibria—and the various sorts of noise present in actual systems may prohibit in principle the sort of perfect information that separating equilibria require. Separating equilibria require a system without sender error, dishonesty, polysemy, ambiguity, etc. Given that it seems unlikely that such a system might exist, it seems unlikely that the nearest separating equilibrium solution will work for signaling systems beyond formal abstractions.

Birch’s suggestion for dealing with those systems—that have no separating equilibria in their evolutionary dynamics is that we understand nearness in terms of the changes to the system parameters required to produce a signaling equilibrium. The nearest separating equilibrium for a system that has no separating equilibria in its evolutionary dynamic is the separating equilibrium that would be made possible by changing the evolutionary dynamic as little as possible (Birch 2014, p. 504). That is, the content of a signal in a given signaling system is determined by the behavior that signal at the nearest separating equilibrium of the most evolutionarily similar signaling system that has such an equilibrium. The content of most signals will then be determined by a multi-step abstraction to counterfactuals.

Again, the key concept in determining the appropriate counterfactual system, in this case the ‘smallest parameter change’ is problematic. It is not obvious how we should measure the size of a change to model parameters, especially given that the system may be more sensitive to changes in one parameter than another. We might merely consider the total value of the change, a total weighted by sensitivity, look for the parameter values that produce behavior most like what we have observed, or take any number of other approaches. Birch suggests that we can use the fact that some parameter values are more plausible than others to develop a sense of what changes are permissible, and that this would help us determine what the smallest parameter change should be (Birch 2014, p. 507). However, this simply won’t work because the plausible values for any one parameter in a system depend in some part on the values of other parameters of the system. The alignment of player interests, the cost of signaling or acting, and the value of any number of other parameters are presumably selected for in a way that is sensitive to the communicative properties of the signaling
convention as a whole, which in turn results from those values. One signaling system might evolve low cost signaling that allows unreliable or frivolous communication at a rate that would be implausible for a system characterized by high cost signaling. Raising costs may be implausible for the first, given how readily it signals. Raising error may be implausible for the second, given how high a cost is paid. Which parameter shifts are plausible for a given system depends on the properties of that system, which are themselves products of the parameters that we are shifting. Further, when the parameter settings needed to produce separating equilibria appear to be impossible for any real system (e.g., no misrepresentation, etc.) these should be ruled out as totally implausible values, which brings us back to our original problem.

The Skyrms/Birch approach to psychosemantics faces many challenges. At root, the problems result from a failure of the core feature of Skyrms’s model to resolve the disjunction and misrepresentation problems. This, in turn, results in Birch developing a proposal to resolve them, but this solution hinges on appealing to counterfactuals without making clear to which counterfactuals we should appeal. As a result, the Skyrms/Birch approach falls well short of an acceptable theory of content.

2.2 Shea/Godfrey-Smith/Cao information/function semantics

Another noteworthy causal theory of content is given by Shea, Godfrey-Smith, and Cao (2018), and includes two distinct notions of content. The first is informational content, which is a simple probabilistic reinterpretation of Skyrms’s model. The second is functional content, which is meant to capture the role that a signal plays in stabilizing a system at equilibrium.

Shea et al. define the informational content of a signal as a vector of the post-signaling probabilities of the events about which the system communicates. The simple example of Eq. 3 then has as its content the vector of probabilities 1, 0 instead of the information vector 1, −∞ given by Skyrms. Acknowledging that we can make such a probability vector even when pre and post-signaling probabilities are identical, Shea, et al. stipulate that where probabilities are unchanged, the signal has no informational content (Shea et al. 2018, pp. 1013–1014). This approach to informational content is effectively Skyrms’s model cast in probabilistic terms. As a result, the same problems that prevent Skyrms’s model from providing an adequate theory of content arise here. The seriousness of these problems has already been addressed, so we need not say anymore about Shea et al.’s notion of informational content.

However, Shea et al.’s model of functional content requires further consideration. Functional content is a product of the role that a signal plays in stabilizing a signaling convention at evolutionary equilibrium. The process of identifying the functional content of a signal involves multiple steps. First, we check to make sure that the system is at an evolutionary equilibrium. If it is, then we look to see if there is some behavior that the receiver produces specific to that signal such that the receiver’s behavior given that signal is distinct from its behavior given other signals in the system. If such a behavior exists, we look at whether there is some state of the
world that obtains on at least some occasions when that signal is sent, and where the receiver’s performance of the action specific to the signal given that state of the world contributes to the evolutionary stability of the signaling convention. If such an event exists, this event is the content of the signal. If more than one such event exists, then the signal has disjunctive content. If the signaling system is not at equilibrium, or if there is no such event, or no such behavior, then the signal has no functional content (Shea et al. 2018, pp. 1011, 1014–1016).

Like informational content, functional content is described as a vector. However, the functional content vector of a signal describes the role that each event plays in stabilizing the use of that signal in the convention. Given a signal $s_1$, which for each event $e_i$ in $E$ results in a marginal payoff of $x_i$ which records the advantage gained by signaling and $\sigma$ which is total marginal payoff produced by $s_1$ for all events combined, we can describe the functional content of a signal as follows:

$$F_E(s_1) = \left(\frac{x_1}{\sigma}, \frac{x_2}{\sigma}, \ldots, \frac{x_n}{\sigma}\right)$$

Which says that the functional content of $s_1$ is a vector recording the proportion of total marginal fitness that $s_1$ contributes in case of each event (Shea et al. 2018, pp. 1030–1031).

Conceptually, the functional content vector of a signal records the relative contribution of each event in continuing the use of that signal. This is a product of the benefits gained by using that signal for each state. Accordingly, the functional content vector describes the proportional benefit that is provided by the signal in case of each event. As this benefit is ultimately responsible for the continued use of the signal, this means that the content of a signal is the set of states under which that signal provides an advantage, weighted in accordance to how much of an advantage is provided.

One concern we might have about this model of functional content is that it seems to fall prey to the misrepresentation and disjunction problems under certain conditions. For instance, imagine the signals ‘wolf’ and ‘coyote’. It is presumably incorrect for ‘wolf’ to be sent on the event of seeing a coyote, or for ‘coyote’ to be sent on the event of seeing a wolf. However, the behaviors that are beneficial in dealing with wolves likely overlap somewhat with those that might be beneficial in dealing with coyotes. In other words, crying ‘wolf’ when a coyote is present still contributes some marginal fitness to the system in virtue of its preparing the receiver for a terrestrial predator. Of course, it would be better have been told ‘coyote’, as the best strategies for dealing with wolves and the best strategies for dealing with coyotes are likely distinct—but any strategy that works against one might work reasonably well against either.

In this scenario, Shea et al.’s theory of functional content assigns strange content to each signal. ‘Wolf’ means something like ‘mostly wolves, but a little bit coyotes’, while ‘coyote’ means ‘mostly coyotes, but a little bit wolves’. Even at first glance,

---

4 The calculation of $x_i$ is a multi-step process that would require more space than can be given to it here. Readers interested in better understanding Shea et al.’s model of functional content should refer to appendix 2 of Shea et al. (2018).
this seems wrong. However, the more troubling implication is that ‘wolf’ sent on the event of seeing a coyote is no misrepresentation, nor is ‘coyote’ on the event of seeing a wolf. While the steps involved are a bit more complicated, we end up facing in Shea et al.’s model of functional content the same challenges with the misrepresentation and disjunction problems that we encountered in the above accounts of informational content.

Another issue arises from the way that Shea et al. make use of equilibria. Shea et al.’s proposal resembles Birch’s in that they both depend on equilibria, but there is a difference in the details of how they make use of these equilibria. This difference, it turns out, is extremely significant. Birch’s solution runs into problems through its dependence on poorly defined standards for identifying relevant counterfactuals. Shea et al. do not depend on counterfactuals, and so do not fall victim to these problems. Instead, they restrict their analysis to systems actually at equilibrium (Birch 2014, pp. 501–502; Shea et al. 2018, pp. 1014–1016). As such, they need not worry about referencing counterfactuals. Instead, the problem for Shea et al. is that their insistence upon actual equilibria limits the applicability of their model. Strict equilibrium is rare in the natural world, and we will not often find it beyond the realm of mathematical abstraction. Because of this, we cannot expect that Shea et al.’s model will give us much insight into actual signaling systems. Their model of functional content may be well defined, but this is achieved at the cost of disconnection from the phenomena that we set out to explain—the semantic content of actual signals.

Shea et al. may not be too worried by the failings of this particular model of functional content, as they are primarily interested in defending the relevance of the role of that a signal plays in producing payoffs that stabilize equilibria to its semantic content (Shea et al. 2018, p. 1030). That this particular model fails is not be a strike against the all such approaches, so long as we can demonstrate that this sort of functional content has something to add for our understanding of the content of signaling systems.

However, any equilibrium based model will need to either define a way to identify relevant counterfactual equilibria, or restrict itself to systems at actual equilibria. The challenges faced by Birch have shown that the first path is extremely difficult, while the second is a non-starter for any theory of content that will apply to actual signaling systems full of actual imperfections. Of course, these problems might be overcome, but we would need to be shown how. Further, it is unclear that the notion of functional content has anything to contribute to our psychosemantics. The mere possibility of formulating such a notion is not sufficient to demonstrate that it has any value—we need to be shown how such a notion would improve our understanding of the behavior of signaling systems. Without such a demonstration, it is premature to take Shea et al. to demonstrate the need for, or value of, equilibrium based approaches to content.

At the same time, it is not clear that what Shea et al. are talking about qualifies as content in the first place. Shea et al. acknowledge that informational and functional content may conflict with one another (Shea et al. 2018, p. 1017). This is unsurprising given that they are different concepts, but this yields a natural question of whether they are both semantic concepts—or, at least, semantic concepts of the same kind.

It is possible, of course, that a pluralistic model of content in psychosemantics might be beneficial. However, such pluralism needs to confirm that its plural
proposals in fact address the same underlying phenomena—or at least that they both deserve to be considered content. It is not clear that what Shea et al. identify in their treatment of functional content is, in fact, content of any kind. Certainly, we can talk about the marginal selective advantage produced by a trait without taking that to imply content—in fact we must be able to do this else all selected traits seem likely to become ‘contentful’. Even considering the evolution of signaling systems alone, it is entirely possible that these sorts of evolutionary properties are just not semantic properties. This is not to say that evolutionary properties are unimportant, but that proponents of functional content need to demonstrate that what they are describing is a genuine semantic (or proto-semantic) property.

3 The greatest surprise reduction theory of content

So far, we’ve reviewed two recent causal theories of content, both of which make use of the intuition that the content of a signal is somehow related to the way that it influences the probabilities of various events. I have also shown that both of these theories fail in familiar ways by failing to account for the misrepresentation and disjunction problems, compounded somewhat by issues raised by their use of evolutionary equilibria to provide a more suitable model.

This much has served to establish the seriousness of the misrepresentation and disjunction problems even for cutting edge causal theories of content. In this section, I aim to solve these problems. I will do so by developing a novel approach—the greatest surprise reduction (GSR) theory of content. The GSR retains the basic probabilistic intuition of information semantics while resolving the misrepresentation and disjunction problems. At the same time, it provides some auxiliary benefits that I will discuss at the end of this section.

3.1 The GSR framework

The basic intuition underlying GSR is that a signal’s content is related to the fact that it makes some events less surprising than they would be otherwise. All else being equal, being told ‘it is raining’ makes us less surprised to find that it is raining. Information theory provides a fairly standard tool for measuring how surprising a given event is, called ‘surprisal’\(^5\) which has been important to information semantics from the beginning (Dretske 1981, pp. 9–10, 52). The surprisal of any event \(e\) is calculated as follows:

\[
S(e) = -\log_2(P(e))
\]

This tells us how surprising something is in bits—in virtue of the fact that we are using a logarithm of base two—which is the conventional unit of information in information theory.

\(^5\) Note that surprisal differs from Skyrms’s measure of information carried only by sign.
To see surprisal in action, let’s consider the possible outcomes of rolling a six-sided die (Table 1). What we see is that surprisal intuitions about how surprising an outcome is quite well. More likely outcomes are less surprising, less likely outcomes are more surprising. Outcomes that are certain to occur are totally unsurprising, and outcomes that are impossible are impossibly (i.e., infinitely) surprising.

One important intuition is that we should be concerned with the way that a signal shifts the surprise of each event. The difference between how surprising an event is pre-signaling and post-signaling is its relative surprise, which can be formalized fairly easily. Let \( P(e_1 | s_1) \) denote the conditional probability of the event \( e_1 \) given a signal \( s_1 \), and \( P(e_1) \) denote the mean probability of the event \( e_1 \) across all signals in the target system, then we can determine the relative surprise \( R(e_1, s_1) \), as follows:

\[
R(e_1, s_1) = - \log_2(P(e_1 | s_1)) + \log_2(P(e_1))
\]

The resulting measure of relative surprise gives us a clear picture of the way that a signal changes expectations about the world.

Relative surprise gives us insight into the consequences of a signals for the surprise of an event, but this is not enough to give us propositional content. The content of a signal cannot just be whatever events it shifts the surprisal of, because this would include events made more surprising. Nor can the content of a signal simply be whatever events it makes less surprising, as we might reasonably be less surprised to see that it is snowing after having been told ‘it is raining’—as it is fairly easy to mistake the two at a glance. Failing to account for these considerations would once again open us up to the misrepresentation and disjunction problems.

The key insight that allows GSR to avoid the misrepresentation and disjunction problems is that the content of a signal should be the event for which the signal causes the largest decrease in surprise—or, probabilistically, the event for which the signal causes the greatest proportional increase in probability. That is, the propositional content of a signal is the event that exhibits the greatest surprise reduction when that signal is sent, the event with the minimum relative surprise given that signal, so long as that relative surprise is negative. Finding the content of a signal \( s_j \) requires that we consider for all events \( e \) that are a member of the set of possible events \( E \) the following function:

---

**Table 1** Six-sided die outcome probability/surprisal

|       | Probability | Surprisal |
|-------|-------------|-----------|
| 1–6   | 1           | 0         |
| 1–5   | 0.83        | 0.27      |
| 2,4,6 | 0.5         | 1         |
| 1     | 0.17        | 2.56      |
| 7     | 0           | \( \infty \) |

---

6. Because \( \log(a/b) = \log(a) - \log(b) \) the difference in surprise is related to the proportional increase in probability, and so relative surprisal is the inverse of Skyrms’s measure of information content.

7. A non-negative minimum relative surprise would mean that no events are more probable given the signal than they would be otherwise, and such a signal would have no content.
The propositional content of a signal is the event with the lowest (negative) relative surprise.

The advantages of GSR can be demonstrated by the considering the following simple convention—which I’ll call the ‘barnyard convention’. The barnyard convention involves a sender sending one of six signals (A, B, C, D, E, and F) in response to the event that one or more various animals (Horse, Cow, Duck, Chicken) are observed. The conditional probability of each event given each signal is given in Table 2.

Given the conditional probabilities of each event given each signal, we are able to calculate the average probability of the event across all available signals. Note that this may not be equivalent to the unconditional probability of the event, as the use of a given convention may itself carry information about probabilities of various events. For instance, the use of semaphore—a system for communicating via the positioning of handheld flags—is by itself probably sufficient to indicate a higher likelihood of some events (e.g., ship) and a lower likelihood of others (e.g., mountain), in virtue of semaphore being used almost exclusively at sea.

With the probability of each event conditional on each signal and the mean probability of each event across the signaling system as a whole, we can calculate relative surprise as described by equation six. Using these values for relative surprise, we can then find the content for each signal by applying formula 7 (Table 3).

### Table 2  Conditional probabilities in the barnyard convention

|       | A    | B    | C    | D    | E    | F    | Mean |
|-------|------|------|------|------|------|------|------|
| Horse | 0.7  | 0.28 | 0.01 | 0.01 | 0.3  | 0.02 | 0.22 |
| Cow   | 0.28 | 0.7  | 0.01 | 0.01 | 0.3  | 0.02 | 0.22 |
| Duck  | 0.01 | 0.01 | 0.9  | 0.68 | 0.3  | 0.76 | 0.44 |
| Chicken | 0.01 | 0.01 | 0.08 | 0.3  | 0.1  | 0.2  | 0.12 |

Values in the ‘Mean’ column are rounded to two decimal places here, which affects the calculations of relative surprisal in the following chart. I use the unrounded values to generate Table 3

\[
C(s_1) = \begin{cases} 
\min_{e \in E} (-\log_2(P(e|s_1)) + \log_2(P(e))), & \text{if } \min_{e \in E} < 0 \\
\text{no content, otherwise} & \end{cases} \tag{7}
\]

The propositional content of a signal is the event with the lowest (negative) relative surprise.

### Table 3  Barnyard surprisals and the greatest surprise reduction

|       | A    | B    | C    | D    | E    | F    |
|-------|------|------|------|------|------|------|
| Horse | –1.67| –0.35| 4.46 | 4.46 | –0.45| 3.46 |
| Cow   | –0.35| –1.67| 4.46 | 4.46 | –0.45| 3.46 |
| Duck  | 5.47 | 5.47 | –1.02| –0.62| 0.56 | –0.78|
| Chicken | 3.54 | 3.54 | 0.54 | –1.36| 0.22 | –0.78|

Bolded values indicate the greatest surprise reduction brought on by each signal, and so the propositional content of the signal. Note that the GSR assigns content reliably using only the conditional probabilities of each event given each signal and the mean probability of each event across the system as a whole.
3.2 GSR, misrepresentation, and disjunction

The barnyard convention not only demonstrates the basic function of the GSR, but also that it can resolve both the misrepresentation and disjunction problems. As these are long-standing problems in psychosemantics which affect nearly all major theories, the ability of the greatest surprise reduction model to resolve them is a considerable merit to this approach.

To solve the misrepresentation problem, a theory must be able identify cases when a signal is used inappropriately. This is a challenge for most causal theories, because even inappropriate uses of a signal are caused by some event. For information semantics, this means that a signal will carry information about any event which occurs alongside the signal. As a result, we need to be able to allow for a signal to carry information about an event that is not part of its content.

The GSR handles this problem quite easily, and all signals in the barnyard convention misrepresent to some extent. By GSR, the content of a signal is the event with the minimum negative relative surprise. Whenever a signal is produced without that event occurring, we have a misrepresentation. For instance, were signal A to be produced when no horse is present, this would be a misrepresentation.

Even reliable misrepresentation, where a signal is more likely to misrepresent than to represent appropriately, is no challenge for the GSR. Signal D, which correlates most strongly with ‘duck’ but nonetheless means ‘chicken’, is an example of such a signal. Being able to handle reliable misrepresentation as it allows for us to describe content even in those cases that, to use Millikan’s phrase, the signal represents correctly ‘just often enough’ to be an effective signal, but not so often that misrepresentation is less likely than representation (Millikan 1989, pp. 288–289).

The GSR also provides a straightforward solution to the disjunction problem. The content of a signal is the event which has its surprise reduced by the greatest amount by that signal. This focus on minimum relative surprise prevents it from falling prey to the failings of other information semantics that assigning all events made more probable by a signal, or even all events not ruled out. To see how we can rule out a proposed disjunction, consider the possibility that signal A means the disjunction ‘horse or cow’. Both horse and cow events are made less surprising by signal A. However, the relative surprise of the disjunction is roughly – 0.91 bits, which is greater than the – 1.43 bits of surprisal for horse alone. Accordingly, we know that the content of the signal is not the disjunction ‘horse or cow’, but simply ‘horse’.

So we can see that the GSR rules out the sort of broad disjunctions that other theories produce. However, there is a second challenge in solving the disjunction problem—while we need to rule out inappropriate disjunctions, we must still allow that some signals may have legitimate disjunctive content (Fodor 1987, p. 102). Given that GSR is restricts content tightly to only the event(s) which exhibit the greatest surprise reduction, we might be worried of running afoul of this concern.

However, this is not a problem for the GSR. The GSR allows legitimate disjunctions to occur whenever multiple events are tied for the minimum relative surprisal, as in the case of signals E and F. Again, the barnyard convention demonstrates many important features of how GSR handles disjunctions. First, note
that three events—horse, cow, and duck—have the same post-signaling probability given signal E. However, only horse and cow are included in the content of the signal, because content is a matter of shifts in surprise, not post-signaling probability. Next, consider that the events that are part of the content of signal F—duck and chicken—have different post-signaling probabilities. Again, this highlights that we are concerned with the proportional shift of probabilities, or the relative surprise of an event given a signal, not the post-signaling probabilities themselves. This is a necessary feature if we are to account for reliable misrepresentation as we do above.

Though the GSR technically allows for disjunction, we might reasonably worry that the bar is so high that we won’t find any disjunctions beyond simple toy examples like the barnyard convention. For a signal to have as its content a disjunction of multiple events, all of those events must have their surprise reduced by an exactly equal amount. For multiple events to exhibit the exact same surprisal shift in a natural system seems unlikely, and so GSR may allow disjunctions in principle but not in fact.

Luckily, this problem is readily solved by considering the noise and uncertainty inherent to real signaling systems. Real signaling systems will need to include an error term to account for their imprecision, and the result will be transforming relative surprisal from a single value to a range. We can say then that the content of a signal is the event with the lowest lower bound relative surprisal and any events the lower and upper bound relative surprisals of which are fall within the range of the lower and upper bound relative surprisals of that event. This shows that the GSR is only artificially precise for artificially precise signaling systems. Accommodating for the imprecision of natural systems suffices to broaden the GSR’s treatment of content so that it become less restrictive in its treatment of disjunctive content. So long as the imprecision of the system is within reasonable bounds, we will have an adequate solution to the disjunction problem that both prohibits unreasonably broad disjunctive contents and allow for appropriate disjunctive contents.

### 3.3 GSR and illocutionary role

Beyond its ability to define content in a way that resolves both the misrepresentation and disjunction problems, the GSR has a number of peripheral applications. One particular interesting application comes from connecting the GSR to a proposal from Zollman (2011) to make sense of the assertive-directive distinction in terms of information content. The assertive-directive distinction is a matter of the illocutionary role of a signal—the sort of function that signal is meant to perform. Illocutionary role is a key component of the pragmatic content of representations, and so is a central phenomenon to be explained in models of representational processes (Bach 2006; Searle 1976). While illocutionary role has mostly been

---

8 Thanks to an anonymous reviewer for pointing out this challenge to me.

9 The appropriate model of this range will depend largely upon the features of the system being considered and how the noise in that system behaves.
explored in context of linguistic states, recent work has shown a growing interest in how these notions might apply to mental states (e.g., Millikan 1995). However, so long as it is unclear how well our psychosemantics can handle this distinction, we might be suspicious that we are missing something important (Zollman 2011, p. 161).

The classic example of the assertive-directive distinction is Anscombe’s shopping list:

Let us consider a man going round a town with a shopping list in his hand. Now it is clear that the relation of this list to the things he actually buys is one and the same whether his wife gave him the list or it is his own list; and that there is a different relation when a list is made by a detective following him about. [I]f the list and the things that the man actually buys do not agree, and if this and this alone constitutes a mistake, then the mistake is not in the list but in the man’s performance [...]... whereas if the detective’s record and what the man actually buys do not agree, then the mistake is in the record (Anscombe 1963, p. 56 emphasis original)

Anscombe’s shopping list demonstrates that even signals of identical form and composition may differ in their content. The detective and shopper may have the same list, but the content of one is a report of the world as it is, while the content of the other guides action to change the world. The detective’s list is an assertive that asserts a fact about the state of the world observed by the detective, while the shopper’s list is a directive that directs the shopper to act in a specific way (Searle 1976, pp. 1–4, 10–11).

Zollman (2011, pp. 162–166) proposes that we can account for the assertive-directive distinction by looking for an asymmetry in the information that a signal carries about states and acts. As the handling of content in Zollman’s work is largely based on Skyrms, Zollman inherits some of the challenges of that approach. However, the key insight from Zollman can be translated rather easily in terms of the GSR to great effect.

The GSR treatment of the assertive-directive distinction is simple and straightforward. If the set of events included in the content of a signal includes only states, then it is an assertive signal. If they include only acts, then it is a directive signal. If they include both, then it is a sort of hybrid signal—a category which is common in animal signaling, and may include performatives as well (Millikan 1995, p. 186). This is, at the very least, an interesting opportunity to make sense of some elements of illocutionary role within the context of a naturalistic psychosemantics.

3.4 GSR and predictive coding

Another interesting consequence of the GSR is that it connects nicely with recent work defending the ‘predictive mind hypothesis’. The predictive mind hypothesis is a general theory of mind built around predictive coding, a tool initially developed for compressing signals to enable more efficient communication. The predictive mind hypothesis holds that mental activity is characterized by a constant flow of
predictions generated by the mind itself, with inputs to those processes carrying only the information required to correct those predictions where they err.

While the formal description of predictive coding is somewhat time consuming, and so set aside here, the basic concepts are fairly familiar. Consider describing the new layout of shopping center to someone who has been there before, and so is familiar with the old layout. Rather than trying to exhaustively describe the entire layout of the shopping center, we can instead describe just the differences between the old layout and the new one. For instance, that the old coffee shop has been paved over and turned into parking lot. This short message would be enough to give a complete understanding of the new layout to someone well acquainted with the old one. It would be much more difficult to give the same level of understanding to someone who had no knowledge of the earlier layout.

The predictive mind hypothesis is that minds are perpetually active predictive machines, in the business of predicting the flow of information from the world by producing models of the world and the informational regularities therein. Mental operations across the board are then a product of the downward flow of predictions produced by generative models and the upward flow of error signals that correct those predictions. If the predictive mind hypothesis is correct, then predictive coding is promising as a general model of mental function, with significant consequences for how we think of the mind (Clark 2013, pp. 1–4, 20–21). This hope is supported by recent work suggesting that predictive coding is implicated in perception (Hosoya et al. 2005; Spratling 2012), action planning (Friston et al. 2010), and cognition (Clark 2013; Spratling 2016).

It has been noted that predictive coding is broadly representational. Both the predictions produced by generative models, and the error signals that correct them, appear to satisfy the standard requirements of representation (Gladziejewski 2016; Wiese 2017). However, there is no consensus on how to assign content to these representations—how to move from the mathematical content—an input–output characterization of the behavior of the system—to the cognitive content—the familiar sort of representational aboutness that psychosemantics tries to explain. This challenge, some hold, is overwhelming—while the mathematical contents of generative models might be used to narrow what sort of cognitive content we can ascribe to them, it will not be decisive (Wiese 2017, pp. 723–724).

However, the GSR generalizes unproblematically to a predictive coding context. Determining the semantics of error signals in predictive coding models is quite like the base case of the GSR. Here, \( s_e \) is an error signal and \( P_g(e) \) is the probability that the generative model assigns to \( e \):

\[
C(s_e) = \begin{cases} 
\min_{e \in E} \left( -\log_2(P(e|s_e)) + \log_2(P_g(e)) \right), & \text{if } \min_{e \in E} < 0 \\
\text{no content}, & \text{otherwise}
\end{cases}
\]

The GSR treatment of an error signal is just the same as any other signal, save that the pre-signaling probability of events is determined by the generative model instead of the mean probability across a signaling system.
The content of the predictions produced by the generative models is derived in a similar way. Let $s_g$ be the prediction produced by the generative model, and $P_0(e)$ be the pre-signaling probability of $e$:

$$C(s_g) = \begin{cases} \min_{e \in E} \left(-\log_2(P(e|s_g)) + \log_2(P_0(e))\right), & \text{if } \min_{e \in E} < 0 \\ \text{no content, otherwise} \end{cases}$$

(9)

It is somewhat more complex to determine the content of a prediction than an error signal because it is not immediately clear how we should determine the pre-signaling probabilities represented by $P_0(e)$. I think that the appropriate pre-signaling probability will depend somewhat on the details of the signaling system. In cases where the signaling system has an ingrained bias such that some events are more likely to be indicated than others, this bias might be represented by setting the pre-signaling probabilities of each event in line with that bias. In other cases, when the sensor is unbiased, we might assume indifference with respect to the pre-signaling probabilities of each event.

Establishing the exact terms in which we should understand the pre-signaling probabilities for defining the content of prediction signals will require further work. However, there is significant promise that GSR may serve as a psychosemantics for the proponents of the predictive mind hypothesis. This would have significant advantages for both GSR and the predictive mind hypothesis. In general, a good psychosemantics benefits from a good theory of mind by allowing it to describe how content is instantiated and functions. Accordingly, the GSR benefits from the predictive mind hypothesis as we can now place the instantiation of mental content in this larger context. At the same time, a good theory of mind benefits from a good psychosemantics by allowing it to explain how states of the mind gain content, and how that content might be used. Accordingly, the predictive mind hypothesis benefits from the GSR by being able to make sense of intentional properties of the mind. Of course, there is a great deal more work to be done in exploring the interactions of the GSR and the predictive mind hypothesis, but it is at present a promising project.

4 Concluding remarks

In this paper, I have shown that the GSR provides an information theoretic solution to some of the most persistent and troubling challenges of naturalistic psychosemantics, promises to demystify some aspects of signal meaning that are often difficult to explain, and may help unify psychosemantics with a larger theory of mind. Because of this, I believe that the GSR provides an important foundation for moving forward in our understanding of the mind. Of course, there is a great deal more work to be done, but I believe nonetheless that the GSR marks a significant advance for psychosemantics writ large.

Before closing on this optimistic conclusion, it is worth taking some time to review one significant problem which remains to be addressed. Birch shows that game-theoretic and information-theoretic approaches to content run into a problem.
in that they assume a way of dividing the world into events. However, there is a considerable amount of flexibility in how we might divide events up, and so this assumption may end up being problematic. This is what Birch calls the ‘partition problem’ (2014, pp. 506–509).

Not only is the partition problem a serious threat, it is actually quite a bit more severe than Birch describes. Birch considers only the problems faced in partitioning the space of states of the world, but similar problems apply to the partitioning of the signal space and the action space on any signaling system. Just as there is not always an obvious way to distinguish one state from another, there are not always obvious ways to distinguish one signal or act from another. Given the fact that states, signals, and acts are meant to connect to one another in a meaningful way, the result is a function containing three parameters the potential values of which are all free to vary considerably. This result of this is that for any way of carving up the system that we choose, there will be a great many others which we might have chosen just as well.

Making the matter even worse, the partition problem extends beyond game-theoretic and information-theoretic approaches to content. Teleosemantic theories also seem to fall prey to a form of the same problem. In fact, we can take the debate between ‘high church’ and ‘low church’ teleosemantics\(^{10}\) to be one largely of how finely we partition states. It is no more obvious for the teleosemanticist how we should partition states, signals, and acts than it is for game-theoretic and information-theoretic approaches to content.

The GSR is straightforwardly vulnerable to the partition problem. It is, perhaps, a small comfort that this vulnerability is shared by nearly all other prominent psychosemantic theories, but a comfort is not a solution. However, I believe that this wide vulnerability to the partition problem in psychosemantics is due to the fact that the partition problem is fundamentally distinct from the problems of content that form the core questions of psychosemantics. We shouldn’t expect to learn how to partition the signaling system from a theory of signal content any more than should expect to learn how to distinguish various possible outcomes from one another by referencing a theory of probability. Resolving the partition problem will be an important step in moving forward with the naturalization of the mind, but it is not the domain of a theory of content.

However, I do believe there is some indication of how we might solve the partition problem. Here, there is something to Birch’s suggestion that the payoffs of various events might be helpful (2014, p. 509). To go a bit further, I believe that the answer to the partition problem will involve understanding the way that the system was organized. For natural systems this will mean invoking an evolutionary framework, in which payoffs will play a critical role. I do not know how closely this will resemble evolutionary stories told elsewhere in the theory of mind (e.g., teleosemantics, Shea et al.’s functional content, etc.). However, I believe that it is here, in understanding the origin and structure of signaling systems, not in the project of defining signal content, that the evolutionary story will be indispensable.

\(^{10}\) See Neander (1995) for a review of this debate.
Addressing the partition problem will be a substantial undertaking of its own, and one that must be taken up as part of the project of naturalizing the mind. So long as it remains unsolved, our overarching understanding of the mind remains inadequate. Even so, the GSR resolves one of the most enduring challenges of psychosemantics—it provides a theory of content that is immune to the misrepresentation and disjunction problems. This, I believe, is a significant contribution.

Acknowledgements This paper began as part of my dissertation, and I owe a great debt to the comments provided to me by my advisor—Jesse Prinz—and my committee members—Barbara Montero and David Papineau. I would also like to thank an anonymous reviewer at this journal whose comments enabled me to strengthen the paper a great deal.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit https://creativecommons.org/licenses/by/4.0/.

References

Anscombe, G. E. M. (1963). Intention (2nd ed.). Cambridge, MA: Harvard University Press.
Bach, K. (2006). Speech acts and pragmatics. In M. Devitt & R. Hanley (Eds.), The Blackwell guide to the philosophy of language (pp. 147–167). Malden, MA: Wiley-Blackwell.
Birch, J. (2014). Propositional content in signalling systems. Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition, 171(3), 493–512. https://doi.org/10.1007/s11098-014-0280-5.
Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behavioral and Brain Sciences, 36(3), 181–204. https://doi.org/10.1017/S0140525X12000477.
Dretske, F. (1981). Knowledge and the flow of information. Cambridge, MA: The MIT Press.
Fodor, J. A. (1987). Psychosemantics: The problem of meaning in the philosophy of mind. Cambridge, MA: The MIT Press.
Friston, K. J., Daunizeau, J., Kilner, J., & Kiebel, S. J. (2010). Action and behavior: A free-energy formulation. Biological Cybernetics, 102(3), 227–260. https://doi.org/10.1007/s00422-010-0364-z.
Gladziejewski, P. (2016). Predictive coding and representationalism. Synthese, 193(2), 559–582. https://doi.org/10.1007/s11229-015-0762-9.
Hosoya, T., Baccus, S. A., & Meister, M. (2005). Dynamic predictive coding by the retina. Nature, 436(7047), 71–77. https://doi.org/10.1038/nature03689.
Millikan, R. G. (1989). Biosemantics. Journal of Philosophy, 86(6), 281–297.
Millikan, R. G. (1995). Pushmi-pullyu representations. Philosophical Perspectives, 9(A1), 185–200.
Neander, K. (1995). Misrepresenting & malfunctioning. Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition, 79(2), 109–141.
Rupert, R. D. (2008). Causal theories of mental content. Philosophy Compass, 3(2), 353–380. https://doi.org/10.1111/j.1747-9991.2008.00130.x.
Searle, J. R. (1976). A classification of illocutionary acts. Language in Society, 5(1), 1–23.
Shea, N., Godfrey-Smith, P., & Cao, R. (2018). Content in simple signalling systems. The British Journal for the Philosophy of Science, 69, 1009–1035. https://doi.org/10.1093/bjps/axw036.
Skyrms, B. (2010a). Signals: Evolution, learning, & information. New York, NY: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199580828.001.0001.
Skyrms, B. (2010b). The flow of information in signaling games. *Philosophical Studies, 147*(1), 155–165. https://doi.org/10.1007/s11098-009-9452-0.

Skyrms, B., & Barrett, J. A. (2019). Propositional content in signals. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 74*(April), 34–39. https://doi.org/10.1016/j.shpsc.2019.01.005.

Spratling, M. W. (2012). Predictive coding as a model of the V1 saliency map hypothesis. *Neural Networks, 26*, 7–28. https://doi.org/10.1016/j.neunet.2011.10.002.

Spratling, M. W. (2016). Predictive coding as a model of cognition. *Cognitive Processing, 17*(3), 279–305. https://doi.org/10.1007/s10339-016-0765-6.

Wiese, W. (2017). What are the contents of representations in predictive processing? *Phenomenology and the Cognitive Sciences, 16*(4), 715–736. https://doi.org/10.1007/s10339-016-0765-6.

Zollman, K. J. S. (2011). Separating directives and assertions using simple signaling games. *The Journal of Philosophy, 108*(3), 158–169.

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.