Modality, scale structure, and scalar reasoning

Daniel Lassiter
Stanford University

Abstract Epistemic and deontic comparatives differ in how they interact with disjunction. I argue that this difference provides a compelling empirical argument against the semantics of Kratzer (1991), which predicts that all modal comparatives should interact with disjunction in the same way. Interestingly, an identical distinction is found in the semantics of non-modal adjectives: *additive* adjectives like 'heavy' behave logically like epistemic comparatives, and *intermediate* adjectives like 'hot' behave like deontic comparatives. I characterize this distinction formally and argue that the divergence between epistemic and deontic modals explained if we structure their semantics around scalar concepts: epistemic modals should be analyzed using probability (an additive scale), and deontic modals using expected value (an intermediate scale).

A prominent line of research takes it both possible and desirable to provide a unified semantics for deontic, epistemic, and other modal notions (Kratzer 1977, 1981, 1991, 2012). In the first part of the paper I present a puzzle about variable interaction with disjunction that appears to call the feasibility of this project into question, and then point out a related distinction between two kinds of adjectival scales: additive and intermediate. I'll then suggest a way to maintain an overall unified approach using tools developed for non-modal adjectives: we can adopt recent proposals to treat both kinds of modals as having a semantics build around degree scales rather than (or at least in addition to) quantification over worlds. This discussion has both negative and positive implications. On the negative side, it reveals some empirical problems with received theories of the semantics of modal language. On the positive side, it suggests a way to repair these problems by treating the semantics of scalar modals as a core issue in modal semantics. I will sketch some arguments for the utility of this perspective in understanding both epistemic and deontic modals, and conclude with a methodological remark on the place of decision-theoretic, probabilistic, and other scalar structure in the lexical semantics of modals and other expressions.

1 The core puzzle: *Likely* and *Good*

Six people — Alice, Bob, Charlie, Doug, Elizabeth, and Francis — are playing a high-stakes game. Each of them is assigned a different number between 1 and 6. A single die will be thrown at noon tomorrow. Minutes later, a $100,000,000 prize, provided by an anonymous benefactor whose motivations remain mysterious, will be deposited into the bank account of the person who has been assigned the number that comes up on the die.

We can imagine the story continuing in several different ways, crossing the likelihood of the various possible outcomes with their moral desirability.

- Likelihood manipulation:

- **Equal likelihood**: The die is unweighted: each person is equally likely to win.
- **Unequal likelihood**: The die is heavily weighted toward one outcome (say, the number 1). Alice has been assigned the number 1, so she is much more likely to win than anyone else.

• Goodness manipulation:

- **Equal goodness**: The characters are indistinguishable in motivations and likely actions: all of them will behave in pretty much the same way (good or bad) if they win.
- **Unequal goodness**: One character is much more virtuous than the others: say, if Alice wins she will dedicate the winnings to cancer research, feeding the poor, and a constitutional amendment to overturn the *Citizens United* decision, while the others would fritter it away on yachts and caviar.

Suppose first that we are in an **Unequal likelihood** condition (of either Goodness type). Clearly A = *Alice wins* is more likely than B = *Bob wins*, and A is also more likely than C = *Charlie wins*. If the die is very heavily weighted, it might even be that A is more likely than $B \vee C$. Since *more likely than* entails *at least as likely as*, we can conclude that the following sentences are **jointly satisfiable** (making the obvious adjustments for English grammar and word order in the unabbreviated forms).

(1) a. A is at least as likely as B.
b. A is at least as likely as C.
c. A is at least as likely as $(B \vee C)$.

Although (1c) could well be true in the scenario described, it is clear that it is not a valid inference from (1a) and (1b). To see this, simply switch to the **Equal likelihood** condition. Here, A, B, and C are all equally likely, but $B \vee C$ is rather more likely than A — exactly twice as likely, in fact. This means that (2) is **not valid**.

(2) INVALID INFERENCE:
a. ϕ is at least as likely as ψ.
b. ψ is at least as likely as χ.
c. $\therefore$ ϕ is at least as likely as $(\psi \vee \chi)$.

If this were valid, we could feed the conclusion back in as a premise, using the fact that A is pairwise as likely as D, E, or F to derive an even more preposterous conclusion: *A is as at least as likely as* $(((B \vee C) \vee D) \vee E) \vee F$. Since one and only one person will win, this is equivalent to *A is at least as likely as* $\neg A$, which is plainly false in the **Equal likelihood** condition: Alice is much more likely not to win than she is to win.

The situation is rather different when we replace *likely* with *(morally) good*. In the **Unequal goodness** condition (of either Likelihood type), Alice's winning would be a moral good since she would use the money to promote good causes, while the others would use their winnings for things which are morally indifferent or worse. Clearly, then, A is pairwise better than B, C, D, E, and F;

so, of course, A is at least as good as each of these as well. In this case, all of the following are clearly true:

(3) a. A is as at least as good as B.
   b. A is as at least as good as C.
   c. A is as at least as good as $(B \vee C)$.

It is at least as good if Alice wins as it is if Bob or Charlie wins, in fact much better: if one of them wins, this is at best a morally indifferent outcome while Alice's winning will lead to great moral good.

In the case of *likely*, we learned that the analogue of (3c) was not a valid inference by switching to the **Equal likelihood** condition. Strikingly, an analogous move here does not affect the situation. If we switch to the **Equal goodness** condition, it is completely morally indifferent which of the possible outcomes emerges. It seems equally clear that each is exactly as good as the disjunction of the others, and that each is exactly as good as its own negation. All of the sentences in (3) are true here, too.

Manipulating the likelihood of the various outcomes affects joint satisfiability of sentences of this form involving *likely*, while manipulating the goodness of outcomes in equivalent ways apparently does not affect otherwise identical examples involving *good*. This follows if (2) is invalid but (4) is valid, as it intuitively is:

(4) VALID INFERENCE
   a. ϕ is as at least as (morally) good as ψ.
   b. ϕ is as at least as (morally) good as χ.
   c. $\therefore$ ϕ is as at least as (morally) good as $(\psi \vee \chi)$.

If this is right, we can conclude that the likelihood of disjuncts 'adds up' when they are disjoint, while the goodness of disjuncts does not. This has important effects on inferences involving these expressions: the goodness of a disjunction is **at most as great as** the moral goodness of the best disjunct, while the likelihood of a disjunction will frequently be **greater than** the likelihood of any disjunct. From this we can conclude that likelihood and goodness do not have the same logic.

Likely and *good* are the most basic components of English expressions of epistemic and deontic comparison, and they have many other grammatical and logical similarities. Such a sharp divergence in the inferences that they license poses a serious challenge to widely held belief — usually based on observations about epistemic auxiliaries alone — that the logical structures underpinning epistemic and deontic concepts are identical. For example, in Kratzer's (1991) semantics both 'ϕ is as likely as ψ' and 'ϕ is as good as ψ' receive truth-conditions from a binary relation on propositions ('Comparative Possibility') with an identical logical structure both cases, with the only variation being the specific content supplied by some contextually variable sets of propositions. The Comparative Possibility relation plays a crucial role in accounting for the meanings of several other modal expressions and for the logical relations between modal adjectives and (e.g.) modal auxiliaries. On a Grand Unified Theory of epistemic and deontic modality like this, (2) and (4) must perforce pattern together — both valid or both invalid. (Kratzer's (1991) proposal makes them both valid, but this is inessential; the point is that a completely unified semantics will not be able to distinguish them.) Such a theory is desirable *a priori*, but the logical differences between *likely* and *good* raise considerable doubt about whether it can be sustained.

2 Background

The choice to focus on the gradable modal adjectives *likely* and *good* in a discussion of epistemic and deontic modality is a bit unusual.And we will consider only their use as propositional operators, ignoring locutions such as *a good car* and *a likely suspect*. The linguistically-oriented literature on modality has dealt almost exclusively with the analysis of auxiliary verbs such as epistemic *must* and *might*, deontic *may, must, should*, and the deontic uses of the quasi-auxiliaries *ought* and *have to*.On the distinction between auxiliaries and quasi-auxiliaries see Huddleston & Pullum 2002: §3. Unfortunately the single most-discussed modal expression, *ought*, is among the quirkiest syntactically. In philosophical logic, modals that are not auxiliary or quasi-auxiliary verbs have played a much greater role, notably the adjectives *necessary, possible, permissible*, adverbs *necessarily, possibly*, and complex expressions such as *have an obligation (that/to)*. *Good* and *bad* have been important players in at some work in deontic logic: see, for example, Chisholm & Sosa 1966; Hansson 1990; Gustafsson 2014; Carlson 2014. The sophisticated discussions of these expressions in philosophical logic have nevertheless left much to be determined in how the proposals can be integrated with the rest of what we know about the syntax and pragmatics of natural languages and their lexical and compositional semantics. I'll take it for granted here that such an integration is a fundamental requirement on a complete theory of modal semantics.

The choice to focus on *likely* and *good* here has several overlapping motivations. Most obviously, there is an important lacuna in the existing literature's descriptive coverage which deserves to be filled. In addition, modal adjectives have clear logical relations to the better-studied auxiliary modals, in both directions. For example, there are intuitively valid implications in both directions between epistemic *must* and comparatives formed with *likely* (5-6), and between deontic *ought* and comparatives formed with *good* (7-8).

(5) a. Mary must be at home by now.
b. ∴ More likely than not, Mary is at home.

(6) a. It's likely that Bill is in Paris.
b. ∴ He might be in Paris.
c. ∴ It's not the case that he must be in New York.

(7) a. Since staying would offend her host, Mary ought to leave.
b. ∴ It's (morally) better if she leaves than if she doesn't.

(8) a. It's (morally) far better if Bill goes to Paris than if he goes to New York.
b. ∴ It's not the case that he ought to go to New York.

Even if (for some odd reason) we were to stipulate that only the modal auxiliaries and quasi-auxiliaries are of theoretical interest, these inferences would be enough to show that a theory of their semantics cannot afford to ignore the evidence provided by the modal adjectives *likely* and *good*.

In fact, I believe that scalar concepts have a fundamental role to play in modal semantics, and a much greater role than has been acknowledged in most of the literature (Lassiter 2011, 2014d,b). If this is correct, an understanding of the semantics of scalar modals such as *good* and *likely* has the potential to illuminate important features of epistemic and deontic concepts which have been obscured by the literature's near-exclusive focus on a small closed class of modal expressions. Let me expand a bit on this latter point.

Likely and *good* are gradable adjectives with the usual range of grammatical behavior: degree modification (*very likely/good*), comparative structures (*better/more likely than*, *as good/likely as*), and so on. Given this, it is natural to suppose that, like other gradable adjectives, they have a semantics built around *degrees* which are organized into *scales* (or one built around a closely related order-theoretic structure: see Lassiter 2014a for a discussion of the relationship between degree-, delineation-, and order-theoretic semantics for scalar adjectives).

Modal verbs are grammatically more limited in various ways than modal adjectives, and in particular some (but not all) modal verbs are restricted in their ability to participate in gradation and comparison structures. So, if many or all deontic concepts are fundamentally comparative and scalar as I am supposing, we are more likely to gather reliable insights into the the formal structure of deontic scales by looking at the class of deontic items which participates most freely in gradation and comparison. Likewise, if scales play an important role in the analysis of epistemic concepts, then it will be difficult to inspect the formal structure of their scales by analyzing modals which do not participate in gradation and comparison.

3 Additive and intermediate adjectives

The core puzzle of §1 — the different patterns of entailment in epistemic and deontic comparisons (2) and (4) — is not particular to modal adjectives. Semantic analyses of adjectives drawing inspiration from Measurement Theory (Krantz, Luce, Suppes & Tversky 1971) have occasionally noted the importance of additivity in the semantics of certain scalar expressions (Krifka 1989; Nerbonne 1995; Sassoon 2010; van Rooij 2011). A typical example is *heavy*: if you and I stand on a scale together, the weight that registers will be the sum of our individual weights. It follows immediately that an object x can be pairwise as heavy as (non-overlapping) objects y and z, but fail to be as heavy as the compound object formed by putting y and z together. As we'll see, plurals and conjunctions of individuals have a close formal connection to disjunctions of propositions. If so, there is a precise parallel between the likelihood of disjunctions and the weight of compound objects: both add up.

For concreteness, I adopt the mereological theory of plurality of Link (1983), where the domain of individuals has a part/whole structure. Various alternatives theories of plurality are available — see Scha & Winter 2014 for a survey — and the choice among them is not very important for current purposes. In Link's theory, for any two count objects a and b, there is a third object $a \sqcup_i b$, read as 'the individual-join' or 'i-join of a and b'. (This operation is also called 'sum formation' or, in Measurement Theory, 'concatenation'). In Link's theory, the domain of individuals forms a join semilattice with no bottom element, as for example in Figure 1. In this figure, an arrow from a node x to a node y indicates that y is a part of x. (Reflexive and transitive arrows are predictable, and are suppressed for readability.)

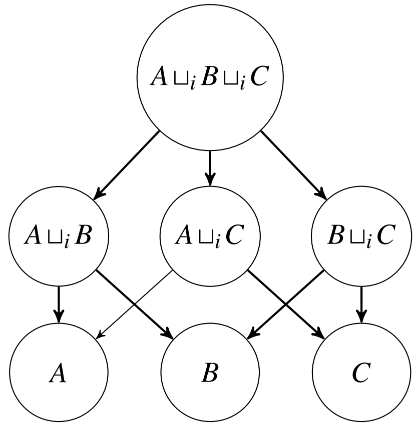

Figure 1 Structured domain of individuals with three atoms A, B, and C.

The (i-)join of two objects x and y is the least node dominating both.More formally: the unique z such that there is an arrow from z to each of x and y, and there is no $z' \neq z$ such that there is an arrow from z to z' and an arrow from z' to each of x and y. Note that the arrows in question may include the predictable transitive and reflexive arrows which are not shown in these figures, to avoid clutter. So, for example, the i-join of a and b is the node labeled $a \sqcup_i b$ — the object consisting of a, b, and nothing else. The i-join of $a \sqcup_i b$ and $b \sqcup_i c$ is $a \sqcup_i b \sqcup_i c$. For any node x, the i-join of x with itself is just x again: $a \sqcup_i a = a$.

Importantly, the i-join operation in the semilattice of count individuals corresponds formally to disjunction $\vee$ in the lattice of propositions, in the sense that disjunction in propositional logic corresponds to the join operation in algebraic logic.See Halmos & Givant 1998 for an excellent exposition of basic algebraic logic. Figures 1 and 2 illustrate the parallelism between these semantic domains, with a structured domain of individuals with 3 atomic individuals in Figure 1 and the lattice of propositions formed from a 3-world domain of possible worlds in figure 2. The only structural difference between them is that the propositional domain has a bottom element $\varnothing$ (representing the denotation of a contradiction). In the propositional domain, an arrow from x to y indicates that y is a subset of x, i.e., that y entails x.

Each property of i-join noted above has corresponds to a property of $\vee$ in the lattice of propositions in figure 2. The join (disjunction) of $\{w_1\}$ and $\{w_2\}$ — $\{w_1\} \vee \{w_2\}$ — is equal to the the least node which dominates both, their union $\{w_1, w_2\}$. The join of $\{w_1, w_2\}$ and $\{w_2, w_3\}$ is $\{w_1, w_2, w_3\}$. The join of an proposition A with itself is A again: e.g., $\{w_1\} \vee \{w_1\} = \{w_1\}$.

Returning to adjectives such as *heavy*, then, we want to know if the mereology just described can help us model the observation that the weight of the plural individual formed of you and me is the sum of our individual weights. We can do this by requiring that the relevant measure function has the following property:

(9) If μ_A is ADDITIVE, then, for all non-overlapping x and y, $\mu_A(x \sqcup_i y) = \mu_A(x) + \mu_A(y)$.

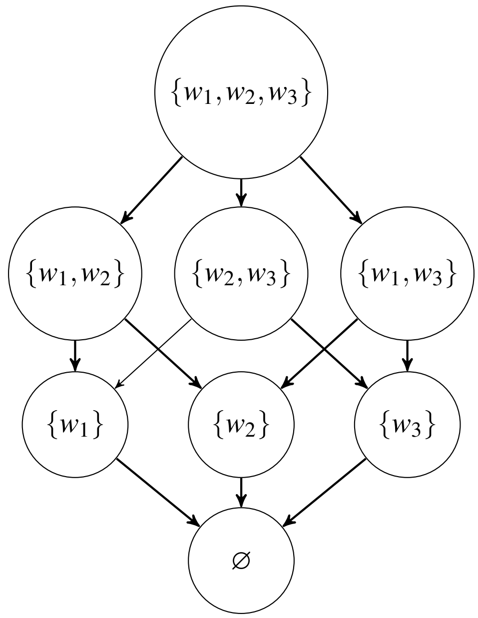

Figure 2 Powerset lattice generated by a 3-world W. If the bottom element $\varnothing$ were removed, this structure would be isomorphic to the semilattice in figure 1.

If μ_A is additive, I will say that A is an ADDITIVE PROPERTY and that A lives on an ADDITIVE SCALE. The restriction to non-overlapping x and y is important: even though my arm has non-zero weight, we do not want to predict that $\mu_{heavy}(\textbf{Dan} \sqcup_i \textbf{Dan's arm})$ is the sum of $\mu_{heavy}(\textbf{Dan})$ and $\mu_{heavy}(\textbf{Dan's arm})$. Rather, since my arm is a proper part of me, $(\textbf{Dan} \sqcup_i \textbf{Dan's arm}) = \textbf{Dan}$. The weights are thus the same.

Suppose Alice (A) weights 140 pounds, and Bob (B) and Charlie (C) weigh 120 pounds each. Then (10a) and (10b) are true, but (10c) is false: $\mu_{heavy}(B \sqcup_i C)$ is 240 pounds, which is greater than Alice's weight.

(10) a. A is as at least as heavy as B.
b. A is as at least as heavy as C.
c. A is as at least as heavy as $(B \sqcup_i C)$.

The following inference pattern is not valid, then.

(11) INVALID INFERENCE
a. x is as at least as heavy as y.
b. x is as at least as heavy as z.
c. $\therefore$ x is as at least as heavy as $(y \sqcup_i z)$.

The invalidity of (11) is obviously similar to the invalidity of the related argument discussed above for *likely*. Degrees of heaviness, like degrees of likelihood, add up over joins (/disjunctions).

Lassiter (2011, 2014d) points out that there is an interesting class of scalar properties which interact rather differently with the *i*-join operation, dubbing these INTERMEDIATE. For intermediate properties, there is a systematic relationship between the degree to which two individuals have the property and the degree to which their join has the property, but this relationship is not well-modeled by addition. For example, if I pour two bowls of water of different temperatures into one container, the volume of the result will be the sum of the volumes of water in the two bowls, but the temperature of the result will not be the sum of the temperatures of the two bowls. Rather, it will be some quantity **intermediate** between the temperatures of the two original bowls.

(12) If μ_A is INTERMEDIATE and $\mu_A(x) > \mu_A(y)$, then $\mu_A(x) > \mu_A(x \sqcup_i y) > \mu_A(y)$.

Danger is another plausible case of an intermediate property. Additivity entails that an object has at least as great a degree of the property as any of its proper parts. (A truck is at least as heavy as its front right wheel, indeed more so: its weight is the sum of the weight of the front right wheel and the weight of the rest of the truck.) Danger is clearly not additive: if it were, California would have to be at least as dangerous as South Central Los Angeles, and strictly more so given that there is non-zero danger in other parts of California. But if danger is intermediate, we can make reasonable predictions like this: if north Metroville is very dangerous and south Metroville is very safe, then the right answer to the question 'How dangerous is Metroville?' is not 'very' or 'extremely', but 'moderately'.

An interesting subtype of intermediate adjectives are the **weighted-average** (WA) adjectives. The signature of a WA-adjective is that there the degree of the property assigned to a compound object is sensitive not only to the degrees to which its parts have the property, but also to an additive measure which gives the relative contribution of each part to the result. The prediction, then, is that a compound formed of two parts which are unequal with respect to the weight function will have a degree of the property which is closer to the degree of the part which receives greater weight.

Temperature is an example of a WA-adjective, with volume (or something typically correlated) as its weight function. If I combine two bowls of water of different temperatures, but bowl A contains more water than bowl B, then the temperature of the result will be closer to A's temperature than it will be to B's temperature, with the divergence being controlled by how much larger A is. Formally, we can define:

(13) If μ_A is a WA-MEASURE then, for all non-overlapping x, y and relevant weight function w,

$$\mu_A(x \sqcup_i y) = \frac{\mu_A(x) \times w(x) + \mu_A(y) \times w(y)}{w(x) + w(y)}.$$

Every WA-measure is an intermediate measure, as long as neither $w(x)$ nor $w(y)$ is equal to 0.

Strikingly, the invalid inference in (10) **is** valid if *heavy* is replaced with an intermediate adjective such as *hot*.

(14) VALID INFERENCE
   a. x is as at least as hot as y.
   b. x is as at least as hot as z.
   c. $\therefore$ x is as at least as hot as $(y \sqcup_i z)$.

If x is pairwise as hot as y and z, then of course x will be as hot as $y \sqcup_i z$, because the temperature of $y \sqcup_i z$ is between that of y and z (if they are different, or equal to both if they are the same). This is intuitively correct. The validity has essentially the same character as the validity with *good* in (3). Degrees of heat, danger, etc., like degrees of goodness, do not add up over join/disjunction.

The difference between the modal adjectives *likely* and *good* in their interaction with disjunction thus has a precise analogue in a little-noticed distinction among different kinds of adjectives. Given that a theory of degree semantics needs to allow for adjectives to encode lexically whether they are on an additive or intermediate scale, we may be able to resolve the problem with the modal adjectives while maintaining an overall uniform semantics for epistemic and deontic modals by appealing to this parameter of variation: epistemic scales are additive, and deontic scales are intermediate.

4 Modal scales

Recently, a number of proposals have been made which — taken together – suggest just this. Yalcin (2005, 2007, 2010); Swanson (2006); Lassiter (2010, 2014b) have suggested that *likely*'s scale is ordinary probability — an additive scale. Jackson (1991); Goble (1996) have proposed taking the scale for *good* to be expected (moral) value, an intermediate scale, and have also argued that *ought* should be treated as an expected-value operator. Lassiter (2011, 2014d) articulates a framework which combines these ideas into a single scalar theory of modality, drawing connections with the semantics of scalar adjectives and with Measurement Theory (Krantz et al. 1971).

4.1 *Likely* and Probability

A **probability measure** is a function P which takes propositions (sets of worlds) to real numbers between 0 and 1, subject to three constraints:

- The domain of P contains W and is closed under negation, disjunction, and conjunction.
- $P(W) = 1$.
- For disjoint $A, B \subseteq W$, $P(A \vee B) = P(A) + P(B)$.

(I assume finite W in order to avoid distracting technicalities, but this assumption could be relaxed without affecting any of the arguments given.)

The **conditional probability** of a proposition A, given a proposition B, is a derived probability measure generated, in effect, by assigning measure 0 to the not-B portion of logical space and renormalizing by dividing by $P(B)$.

$$P(A|B) = \frac{P(A \wedge B)}{P(B)}$$

(This is the ratio definition due to Kolmogorov (1933); taking conditional probability as a primitive would be fine for our purposes, too.)

The formal properties of (conditional) probabilities are extremely similar to those of weights. The primary difference is that probabilities can be scaled in a non-arbitrary way because their domain has a fixed maximum element, W. Weights, in contrast, are measures whose domain is theoretically unbounded. The choice of reference element with respect to which other measures are scaled — the foot, inch, mile, meter, etc. — is thus arbitrary.

I'll briefly sketch of a few of the reasons that have been given in recent work for treating *likely*'s scale as a probability scale. One impetus is the observation that the current standard semantics for epistemic concepts (Kratzer's) predicts that the invalid disjunction-involving arguments that we discussed above should be valid.

(15) INVALID INFERENCE:
   a. ϕ is as at least as likely as ψ.
   b. ψ is as at least as likely as χ.
   c. $\therefore$ ϕ is as at least as likely as $(\psi \vee \chi)$.

Lassiter (2010); Yalcin (2010) point out this problem for the proposal in Kratzer 1991. (A revised proposal in Kratzer 2012, intended to resolve this issue, does not fare better: see Lassiter 2014b.) For reasons that we have already seen, this argument is correctly rendered invalid if *likely* lives on a probability scale. This is a good result, but not, of course, proof that probability is the right scale: there are weaker logics for likelihood which also invalidate this inference and have other desirable features (Holliday & Icard 2013).

Another argument that has been given for this analysis of *likely* involves the acceptability of ratio modifiers of the form *(exactly) n times as Adj as*. If I draw a card at random from a well-shuffled pack, it is exactly thirteen times as likely that the card will be a club as it is that the card will be the Queen of Hearts. This is, of course, true because there are thirteen clubs and just one Queen of Hearts. Here are some examples from the web:

(16) Ratio modifiers with *likely*:
   a. Middle-class children in the 1948 cohort were exactly three times as likely as working-class children to be on a five-year course at the age of 15. (Gray, McPherson & Raffe 2012: 203)
   b. [High school] baseball players are exactly 5 times as likely to play some form of professional baseball than [high school] football players are to play in the NFL.

This observation is important because there appears to be a connection between the acceptability of ratio modifiers with an adjective and the (non-)additivity of the scale. Intuitively, one might think that acceptability in phrases like *exactly N times as ADJ as* should depend on whether *ADJ* can be associated with a numerical scale. However, semanticists have noted recently that certain scales that are readily associated with quantities — such as temperature — are nevertheless less frequent and less intuitively natural with *exactly N times as ADJ as*.

(17) Ratio modifiers with additive adjectives:
   a. The truck is exactly twice as heavy/expensive/large as the car.
   b. Bill is exactly four times as rich/tall/fast as John.

(18) Ratio modifiers with non-additive adjectives:
   a. # The truck is exactly twice as hot/ugly/red as the car.

b. # Bill is exactly four times as handsome/angry/disgusting as John.

Modifying a suggestion from Sassoon (2010); Lassiter (2011, 2014b), we can summarize:

> An adjective can form part of a complex ratio modification structure *(exactly) n times as ... as*, with a precise (non-hyperbolic) interpretation, only if it encodes a measure function whose range has a fixed zero point, such as an additive scale.

This constraint makes good formal sense: as Krantz et al. (1971) prove, additivity forces the existence of a unique zero point across all ways of measuring objects. As a result, the question of whether a is exactly N times as *Adj* as b has a unique answer regardless of the choice of unit of measurement (feet, inches, meters, etc.). In contrast, for scales like temperature the choice of zero point is arbitrary, and two ways of measuring temperature (e.g., Fahrenheit and Celsius) can differ in whether $\mu_{temp}(a) = N \times \mu_{temp}(b)$. If the naturalness of certain degree modification structures is sensitive to this difference, we have an explanation of the contrast between (17) and (18).It is, to be sure, possible to find attestations of *n times as Adj as* — without 'exactly' — with adjectives which are clearly not on additive scales, such as *Bill is three times as handsome as John*. As Lassiter (2014b,d) discusses, these uses tend to have a hyperbolic interpretation ('much more Adj than'), and attestations of *exactly n times as Adj as* with non-additive adjectives are virtually nonexistent. Lassiter cites a few naturally-occurring examples which appear to involve temporary absorption of an adjective to an additive scale: *exactly 4 times as funny* (in number of exploding golf balls), or *exactly 100 times as powerful* (specified as being measured in megatons, i.e., amount of TNT that you would need to create an explosion as large).

Note, by the way, that expected value scales of the type I'll discuss below lack a fixed zero point: unlike weight and probability, and like temperature, they are unique only up to positive linear transformation. Thus I predict that examples of *exactly N times as good as* should be rare-to-nonexistent, and examples of *N times as good as* should be hyperbolic, humorous, or involve absorption to an additive scale. My informal investigations (which involved reading the context surrounding many Google hits for 'times as good as') suggest that this prediction is correct. Admittedly, this rationale does not uniquely pick out additive scales, since there could easily be quantifiable, non-additive scales with a fixed zero point (movie star ratings, for example). However, the fact that a scale is additive would readily explain why exact ratio modifiers are acceptable with adjectives on that scale.

If this generalization is correct, then the data in (16) support the hypothesis that *likely* lives on an additive scale. It so, it is a short step to show that it encodes a probability measure: looking at the definition given above, we only need to ensure that

- a tautology receives the maximum possible likelihood;
- if we can talk about the likelihood of ϕ and of ψ, we can also talk about the likelihood of $\neg\phi, \phi \vee \psi, \phi \wedge \psi$ as well;
- any two propositions can be compared in likelihood.

The first two assumptions do not seem to be very controversial. The third is controversial, and it may well turn out that *likely* is associated with a weaker scale which is not connected, but mimics additivity among comparable propositions. (See Keynes 1921; Holliday & Icard 2013 and many others.) Non-connected scales have been suggested for various non-modal adjectives as well (e.g., Cresswell 1976).

There is much more to be said on the topic, and in particular we have not covered important questions about how these conclusions relate to the semantics of other epistemic adjectives (*possible, certain, probable, plausible, clear, evident* and their antonyms), verbs (*think, guess, believe, know*), and auxiliaries (*must, may, can, might*, and perhaps *should, ought*). This brief discussion is intended mainly to demonstrate that additivity is of considerable empirical and theoretical interest, not only in the semantics of non-modal scalar expressions, but in the semantics of scalar modals as well.

4.2 *Good* and Expected Moral Value

Let V be a function which assigns to each possible world w a real number representing its **moral value**. Intuitively, this is a summary representation of how morally desirable it would be if all of the facts of the world were arranged as they are in w, with no remnant uncertainty. The specific numerical values supplied are of course not meaningful, but comparisons between worlds and the relative sizes of differences between values are. That is, we are supposing that moral value is an interval scale, unique up to positive linear transformation (Krantz et al. 1971). This ensures that it makes sense not only to talk about one world being better than another, but also about *how much better*.

The **expected moral value** $\mathbb{E}_V(A)$ of a proposition A is a weighted average of the values $V(w)$ for each $w \in A$, where the weight of each world is given by the conditional probability that it will be actual if A is true. (Here again I simplify by pretending that the set of possible worlds is finite.)

$$\mathbb{E}_V(A) = \sum_{w \in A} V(w) \times P(\{w\}|A)$$

Recall the definition of a WA-measure in (13): for non-overlapping x and y,

$$\mu_A(x \sqcup_i y) = \frac{\mu_A(x) \times w(x) + \mu_A(y) \times w(y)}{w(x) + w(y)}.$$

Expected moral values satisfy this equation, as long as we

- interpret x and y as propositions instead of individuals,
- trade in individual join ($\sqcup_i$) for propositional join ($\vee$), and
- take the additive weight function w to be $P(\cdot|x \vee y)$.

That is, for non-overlapping $A, B \subseteq W$,

$$\mathbb{E}_V(A \vee B) = \frac{\mathbb{E}_V(A) \times P(A|A \vee B) + \mathbb{E}_V(B) \times P(B|A \vee B)}{P(A|A \vee B) + P(B|A \vee B)}.$$

(See Jeffrey 1965: §5.) So $\mathbb{E}_V(\cdot)$ is a WA-measure. Note that that denominator will always be 1, so the equation simplifies to

(19) $\mathbb{E}_V(A \vee B) = \mathbb{E}_V(A) \times P(A|A \vee B) + \mathbb{E}_V(B) \times P(B|A \vee B)$.

Being a WA-measure, $\mathbb{E}_V(\cdot)$ is also an intermediate measure:

(20) $\mathbb{E}_V(A) > \mathbb{E}_V(B) \Rightarrow \mathbb{E}_V(A) > \mathbb{E}_V(A \vee B) > \mathbb{E}_V(B)$,

assuming that neither A nor B has conditional probability 0, given $A \vee B$. This has the following implication: if the expected value semantics for *good* is correct, *good* is non-monotonic. A disjunction can be **worse** than one of its disjuncts (Goble 1996; Lassiter 2011, 2014d).

By contrast, the standard semantics for deontic comparatives (e.g., Lewis 1973, borrowed with slight modifications by Kratzer (1991)) relies on a comparative goodness relation on propositions which is upward monotonic: if ϕ entails ψ, then ψ is at least as good as ϕ. Upward monotonicity holds here because the comparative goodness relation is constructed from a (reflexive, transitive, connected) binary relation $\succcurlyeq$ on possible worlds, and 'ϕ is at least as good as ψ' is defined as meaning 'ϕ is true in at least one world which is ranked at least as high as all ψ-worlds'.

(21) **Lewis-style comparative goodness**: ϕ is at least as good as ψ iff $\exists w \in \phi \forall w' \in \psi [w \succcurlyeq w']$.

If we start with a moral value function instead of a qualitative ordering, we can still give a faithful quantitative rendering of Lewis' proposal, since the value function determines a comparative order: just define $w \succcurlyeq w'$ iff $V(w) \geq V(w')$ and then employ the lifting in (21). $\succcurlyeq$ will then be a reflexive, transitive, connected binary relation.Value functions clearly contain at least as much information as Lewis' $\succcurlyeq$, but they are more informative: V contains additional quantitative information. Above we stipulated that moral value is an interval scale, and so V is unique up to positive linear transformation. This is standard in decision theory (see, for example, Jeffrey 1965 for a clear statement). As a result some, but not all, arithmetic properties of V are stable across transformations. For example, ratios of values $V(w_1)/V(w_2)$ are not stable, but ratio of intervals $(V(w_1) - V(w_2))/(V(w_3) - V(w_4))$ are. A reflexive, transitive, connected binary relation does not, of course, conceal any comparable quantitative information. See Lassiter 2011, 2014d: §2 for discussion of how these measurement-theoretic facts can be integrated into a compositional semantics, and the empirical predictions that emerge. Here **LG** represents the goodness of propositions, and V the value function on worlds.

(22) **Value-theoretic variant**: $\mathbf{LG}(\phi) \geq \mathbf{LG}(\psi)$ iff $\exists w \in \phi \forall w' \in \psi [V(w) \geq V(w')]$.

As Lewis points out, this semantics predicts that the goodness of a (finite) proposition should be equal to the **maximum** of the moral value of any world in the proposition —

$$\mathbf{LG}(\phi) = \arg\max_{w \in \phi} V(w)$$

— and a (finite) disjunction is exactly as good as the best disjunct.

$$\mathbf{LG}(\bigvee\{A_1, ..., A_n\}) = \arg\max_i \mathbf{LG}(A_i)$$

I'll call this property **Maximality**. As a special case, Lewis-goodness (and any semantics with the Maximality property) has the property that, if A is better than B, then A is exactly as good as $A \vee B$.

(23) $\mathbf{LG}(A) > \mathbf{LG}(B) \Rightarrow \mathbf{LG}(A) = \mathbf{LG}(A \vee B)$.

This is in stark contrast to the intermediate property of expected values (20).

The Lewis-style semantics treats *good* as introducing a new kind of interaction with disjunction/join into the typology of scales: in addition to additive and intermediate scales, we would now have Maximal scales. (As an aside, I do not know of any examples of Maximality among the non-modal adjectives: this scale type may be otherwise unattested.) Kratzer's (1991) method of constructing deontic scales ('Comparative Possibility') is formally identical to Lewis', but the semantics differs somewhat because Kratzer does not assume that the underlying world-order $\succcurlyeq$ is connected. This makes the predictions about disjunction more complicated, but the semantics remains upward monotonic, and it is Maximal in connected (sub-)orders.

Both the expected-value semantics and the Lewis/Kratzer semantics correctly validate the disjunction-involving inference in (3). However, there are a number of other differences that can be used to pry apart their predictions. The most obvious is to examine the difference in predictions about entailment directly: when ϕ entails ψ, can ψ be worse than ϕ? On face, it seems that the answer is clearly 'yes'.

A version of Ross' famous puzzle serves to illustrate. After a nice meal cooked for you by a friend, it would be better for you to wash the dishes than not to wash them. Granting this, would it be better for you to *wash the dishes or break them* than to leave them alone? To my mind, the answer is not clear-cut — it depends on what you would be likely to do if given the choice to wash the dishes or break them. But, on a Maximal or upward monotonic semantics for *good/better*, the answer must be 'yes'. Let **UM** be the property of upward monotonicity:

UM) If ϕ entails ψ, then ψ is at least as good as ϕ.

Then we can prove:

1) It's better to wash the dishes than not to wash the dishes.
2) *Better than* and *at least as good as* are transitive relations.
3) *You wash the dishes* entails *You wash the dishes or break them.* [prop. logic]
4) It's at least as good to (wash the dishes or break them) as it is to wash the dishes. [**UM**,3]
5) It's better to (wash the dishes or break them) than not to wash the dishes. [1,2,4]
6) *You don't (wash the dishes or break them)* entails *You don't wash the dishes*. [prop. logic]
7) It's at least as good not to wash the dishes as it is not to (wash the dishes or break them), i.e., to leave the dishes alone. [**UM**,6]
8) It's better to (wash the dishes or break them) than it is to leave the dishes alone. [2,5,7]

This is a *prima facie* implausible conclusion. I'm quite sure that I could accept (1) but deny (8) — particularly if I had a strong suspicion that, given the choice, you would break the dishes. Of the premises, (1) is true by stipulation, and transitivity (2) seems secure. So we have two choices: we can either reject upward monotonicity, or we can find a way to rationalize the odd conclusion.

In fact there are at least two plausible bullet-biting strategies, involving Gricean considerations (Hare 1967; Wedgwood 2006) or free choice effects (von Fintel 2012). These arguments suggest a way to explain the implausibility of the conclusion in terms of special pragmatic features of disjunction which render (8) infelicitous or unassertable in the scenario under consideration, even though it is necessarily true given premise (1). If so, we could maintain upward monotonicity — and the more specific condition of Maximality — in the face of Ross examples.

However, there are examples with an identical logical structure that aren't amenable to a pragmatic explanation invoking special facts about disjunction. Here is one.This is a chancy multi-agent version of Jackson & Pargetter's (1986) 'Professor Procrastinate' example, modified to use *good* instead of *ought*. (I got the idea of 'Procrastinate with probabilities' from Cariani (2013). Jackson (2014), in this issue, also discusses chancy Procrastinate scenarios.) Structurally, it is closely related to the Miners' Paradox (Regan 1980; Kolodny & MacFarlane 2010), Jackson's (1985) Chicken example, and several other deontic puzzles involving decision-theoretic structure.

I switched to using multi-agent Procrastinate cases after finding in experimental investigations that people's intuitions about single-agent cases interact with prior expectations about weakness of will in a way that is difficult to control. As my experimental participants did, I find the relevant judgments to be much clearer in the multi-agent version, because it is easy to structure the example so that the agent who chooses first has no influence on the agent who chooses second except in the choice itself. This point may connect up in an interesting way with Jackson's (2014) discussion of Procrastinate cases in this volume. Jackson's explanation for the *ought*-involving Procrastinate example makes use of the idea that a possibly weak-willed individual who must make sequential decisions can be thought of on the model of multiple agents who choose serially. The example here could perhaps be thought of as generalizing this feature while removing confounds created by intuitions about strength or weakness of will — an issue which Jackson discusses in his paper. (Thanks to Mark Schroeder for pointing out this connection.)

> **Juliet** is considering whether to feign death by taking the drug that Friar Laurence has offered her. If she does, it will put her in a coma, and she will die unless Friar Laurence administers the antidote exactly 10 hours later. If she takes it and the Friar does administer the antidote, she will succeed in convincing her family of her death and she will be able to live happily ever after with Romeo. If she does not take the drug, she will live a long life without Romeo and will be less happy; this is much better than being dead, though. Unfortunately, the Friar is known for being cruel and capricious, and it is extremely likely (though not totally certain) that he will 'forget' to administer the antidote if she takes the drug.

The possible outcomes are partitioned into $O = \{$*No drug*, *Drug* $\wedge$ *antidote*, *Drug* $\wedge$ *no antidote*$\}$. Given the risk involved, and the fact that Juliet has no control over the Friar's actions, an appropriate judgment seems to be:

(24) It is better if Juliet does not take the drug than it is if she takes the drug.

This can be represented quasi-formally by *No drug* $>_{good}$ *Drug*, where *Drug* is equivalent to (*Drug* $\wedge$ *antidote*) $\vee$ (*Drug* $\wedge$ *no antidote*). At the same time, the scenario invites us to consider (25) true:

(25) The best outcome is that Juliet takes the drug and Friar Laurence administers the andidote.

Assuming that the superlative *best* can be spelled out as *parwise better than each alternative*, this can be represented as

(26) $Drug \wedge antidote >_{good} No\ drug$, and $Drug \wedge antidote >_{good} Drug \wedge no\ antidote$.

From (26) and (24) we have (by transitivity)

(27) $Drug \wedge antidote >_{good} Drug$.

On the assumption that logically equivalent propositions are intersubstitutable in this context, we can unpack *Drug* into a comparison of the form *A is better than* $(A \vee B)$:

(28) $Drug \wedge antidote >_{good} (Drug \wedge antidote) \vee (Drug \wedge no\ antidote)$

This contradicts upward monotonicity, which requires that a disjunction be at least as good as each disjunct. Comparative goodness does not appear be upward monotonic, and so *a fortiori* it does not have the Maximality property. I don't know of a plausible strategy for explaining away the data here, but unless one can be found, we must conclude that the Lewis/Kratzer semantics and other upward monotonic theories are incorrect.von Fintel (2012) claims that the original, *ought*-involving Procrastinate examples involve a context shift, but does not give a detailed argument to this effect. It's possible that there are order effects in the original example, as he suggests, but order does not seem to be relevant for the present case involving *good*. Also, it is implausible that all worlds satisfying $Drug \wedge antidote$ are excluded from consideration in forming the judgment that (24) is true, as this style of explanation would require, since the scenario specifically mentions the fact that *Antidote* is possible given *Drug*.

As an aside, Juliet's predicament also has consequences for the semantics of the better-studied (but no more central) deontic items *ought* and *must*. (29) is widely thought to be a valid inference:

(29) a. ϕ ought to be the case.
 b. $\therefore \phi$ is better than $\neg\phi$.

Another widely accepted principle is that *must* ϕ entails *ought* ϕ. If so, we also have that *must* ϕ entails ϕ *is better than* $\neg\phi$. What we have learned about comparative goodness thus places constraints of the form of a semantics for deontic *must* and *ought*.

This may help to understand how sentence pairs like the following could seem reasonable to some in scenarios like **Juliet** (cf. Jackson & Pargetter 1986, etc.):

(30) a. It ought to be that Juliet does not take the drug.
 b. It ought to be that Juliet takes the drug and the Friar administers the antidote.

These sentences are predicted to be inconsistent by any theory in which (29) is valid and comparative goodness has the Maximality property.Note the awkward wording in (30), which is designed to force *ought* to take scope over the the whole conjunction in (30b). This may also force what Schroeder (2011) calls an 'evaluative' reading of *ought*, as opposed to the 'deliberative' readings that are most natural with *Juliet ought not to take the drug*, etc. Deliberative readings of *ought* may well have an interestingly different logic from the evaluative readings under discussion here. (See Chrisman

2012; Finlay & Snedegar 2014 for arguments that the distinction is unnecessary, though.) On the other hand, a semantics for *ought* built around expected moral value can explain why (30a) and (30b) are both intuitively true, while also validating (29). In particular, it predicts that the sentences in (30) are not only logically compatible, but also *true* in situations in which moral value and uncertainty interact in this way; and it does so without invoking additional mechanisms such as context shifting. See Lassiter 2014d: §5-6 for detailed discussion. (A structurally identical explanation can be given for the well-known Ross examples with *ought*, though as above the dialectic is less clear-cut because of the potential for confounds involving disjunction.)

Where does the Lewis/Kratzer semantics go wrong in the **Juliet** scenario (and, perhaps, in Ross examples as well)? My contention is that a Maximal semantics is unable to give a reasonable account of the way that uncertainty influences the weighing of outcomes in moral reasoning. In particular, Maximality instructs us to look for the goodness of a set of possible outcomes by considering how good the **best possible** outcome in that set is, ignoring the issue of how **probable** it is that we will find ourselves among the best outcomes as opposed to the worst, or somewhere in between (cf. Jackson 1985; Goble 1996; Lassiter 2011, 2014d).

But when we are reasoning about what to do under conditions of uncertainty, it is crucial to pay attention to how uncertainty is distributed among the possible outcomes of our actions. The Juliet scenario is carefully constructed to bring this out: some action available to us might bring about the best possible outcome, but is likely to bring about the worst possible outcome, while another choice is guaranteed to bring about a middling result. We judge, against the predictions of a Maximal semantics, that the safe choice is better than the risky choice. The Miners' Paradox (Regan 1980; Kolodny & MacFarlane 2010) has this structure as well, as does the Choice-of-Medicine puzzle introduced by Jackson (1991) and discussed by Goble (1996).

An alternative, non-Maximal semantics is suggested by Cariani, Kaufmann & Kaufmann (2013) (focusing on *ought*). As they point out, the Lewis semantics and its variants effectively enforce the choice rule **Maximax**.This feature is perhaps less problematic if we follow von Fintel (2012); Dowell (2013) in abandoning any connection between the concept of moral value and the world-orderings used in supplying truth-conditions to deontic modals. Then it is possible to simulate the expected-value theory's predictions about certain puzzles involving *ought* (e.g., the Miners' Paradox) by invoking an ordering source function that assigns to each world a set of propositions describing what agents do when they perform the actions that maximize expected moral value. On this approach, the best worlds — which determine the semantic Maximax — will generally be different from the value-theoretic Maximax. Of course, this kind of ordering source will be just one option among many, and we are in need of a well-articulated pragmatic story about why it should be supplied; otherwise it would seem an extraordinary coincidence that 'context' would consistently provide us with precisely the right ordering source to simulate the predictions of a more restrictive competing theory about what *ought* means.

In any case, this tactic would not help with failures of upward monotonicity such as **Juliet** unless we also follow von Fintel (2012) in assuming additional context-shifting operations. But even if it were able to get all the data to come out right, it seems to me much preferable to have a semantics which makes specific, correct predictions about individual cases on the basis of independently motivated value-orderings, without the need to reverse-engineer ordering sources from the observed judgments. See also section 5, where I discuss the dialectic between restrictive and highly flexible theories in this domain. They suggest instead moving to a semantics which enforces **Maximin**: the goodness of an action is equated with the goodness of the *worst* possible outcome of that action.Note, however, that Cariani et al.'s (2013) core ideas can be implemented in a weaker semantics which permits Maximin among many other decision rules, and which is able to cope with the effects of shifting probability. See Cariani 2014 for the details. Like the von Fintel/proposal discussed in the previous footnote, I think that Cariani's extremely permissive semantics is dispreferred on general methodological grounds absent a strong empirical argument against the expected-value semantics. (Cariani suggests several, but they strike me as unconvincing: see Lassiter 2014c.) While this semantics generates good predictions about the Juliet case and some others in the literature, it is not robust to shifts in probability, as we can see by modifying the Juliet story slightly. Suppose we replace 'It is extremely likely (though not totally certain) that the Friar **will not** administer the antidote' with 'It is extremely likely (though not totally certain) that the Friar **will** administer the antidote'. Then our judgments may shift, depending on how important it is for Juliet to be with Romeo: it may be better for Juliet to take the drug. A Maximin semantics does not allow that merely shifting probabilities could influence judgments here. As long the worst possible outcome (death) is still a live possibility, Maximin continues to predict that it is better for Juliet not to take the drug as long as there is any possibility, however slight, that the Friar will not administer the antidote. (For related criticisms see Carr 2012; Cariani 2014.)

On the other hand, the expected-value semantics renders the sentences in (24) and (25) compatible, and it is designed to be sensitive to shifts in probability. For example, consider the model (or rather class of models) described in Table 1. (I simplify by pretending that no other factors impinge on the value of worlds, so that all worlds in each cell of the partition O have the same value. Remember also that value is an interval scale, so that the numerical quantities chosen are not meaningful: only the relative sizes of the gaps between them matter.) The expected values of the

Outcome	value	$P(\cdot \mid Drug)$
$Drug \wedge antidote$	5	.1
$No\ drug$	0	0
$Drug \wedge no\ antidote$	-10	.9

Table 1 A model which verifies both (24) and (25) on the expected-value semantics.

relevant options are thus

- $\mathbb{E}_V(No\ drug) = 0 \times 1 = 0$
- $\mathbb{E}_V(Drug \wedge antidote) = 5 \times 1 = 5$
- $\mathbb{E}_V(Drug \wedge no\ antidote) = -10 \times 1 = -10$

The best option is of course $Drug \wedge antidote$, which has expected value 5. Using the information in

Table 1 and equation (19) we can calculate the goodness of *Drug*:

$$\begin{aligned}\mathbb{E}_V(Drug) &= \mathbb{E}_V(Drug \wedge antidote) \times P(Drug \wedge antidote|Drug)\\ &+ \mathbb{E}_V(Drug \wedge no\ antidote) \times P(Drug \wedge no\ antidote|Drug)\\ &= 5 \times .1 + -10 \times .9\\ &= -8.95\end{aligned}$$

This is much worse than $\mathbb{E}_V(No\ drug) = 0$, and so *No drug* is better than *Drug*. This is possible on the expected-value semantics, even though the **best** *Drug* worlds are better than the best *No drug* worlds, because this semantics makes smart use of probabilistic information.

If we change the story so that the Friar will probably administer the antidote, we can explain the shift in intuitions without modifying anything about the values assigned to the outcomes. In the

Outcome	value	$P(\cdot\|Drug)$
Drug ∧ *antidote*	5	.9
No drug	0	0
Drug ∧ *no antidote*	-10	.1

Table 2 A model which reverses Table 1's predictions about the relative goodness of *Drug* and *No drug* by shifting probabilities.

model described in Table 2, the expected values of the cells of partition *O* are unchanged, but the expected value of *Drug* shifts dramatically: now we have

$$\begin{aligned}\mathbb{E}_V(Drug) &= \mathbb{E}_V(Drug \wedge antidote) \times P(Drug \wedge antidote|Drug)\\ &+ \mathbb{E}_V(Drug \wedge no\ antidote) \times P(Drug \wedge no\ antidote|Drug)\\ &= 5 \times .9 + -10 \times .1\\ &= 3.5.\end{aligned}$$

The fact that our semantics has this flexibility is a good result, since the data suggest that it is necessary: a shift in probabilities can modify judgments about the relative goodness of *Drug* and *No drug* in precisely this way. Of course, this semantics does not **require** such a shift in response to changes in probability. If being with Romeo were not very important, or if death were worse, then even a very high probability of *Antidote* given *Drug* might not be enough to make *Drug* a good choice.

Neither a semantics which encodes unreasonable optimism, nor one which encodes extreme pessimism, is what we need for reasoning about the relationship between uncertainty and graded deontic concepts. **Juliet** and related scenarios suggest that we need a semantics which pays attention to all of the possible outcomes and makes good use of uncertain information. The expected value approach has this character: outcomes are ranked by taking into account all possibilities that they contain, rather than arbitrarily choosing only the best or the worst to consider. The influence of a possibility on the goodness of an outcome is proportional to the probability that the possibility will be actual if the outcome occurs.

This argument does not, of course, show conclusively that expected moral value is the unique correct way to structure deontic scales. It does, however, show that deontic scales have many qualitative features of expected value, and that they do not have the features predicted by a semantics which encodes Maximality — like Lewis's (1973) and Kratzer's (1991) — or another form of upward monotonicity. The fact that scales with this formal structure are needed independently in the semantics of scalar adjectives suggests that this theory is also parsimonious and explanatory.

5 Restrictiveness and neutrality in lexical semantics

I have argued that some prominent accounts in the literature are too restrictive in various ways, and that an expected-value semantics makes better predictions about the interaction between goodness and information. In this concluding section I want to address an argument of a more philosophical character which, if correct, would support adopting a semantics that is still more expressive. Carr (2012) argues that all of the theories we have considered, including the expected-value theory, are undesirable because they 'build controversial normative assumptions into the semantics of modals'. Crucially, she does not give any empirical arguments to the effect that the no semantics which encodes such assumptions could be correct, but seems to take it as obvious that it would be undesirable to do so.To be clear, Carr gives empirically grounded arguments against two specific proposals, those of Kratzer (1991) and Cariani et al. (2013). These arguments do not, nor do they attempt to, establish that *no* semantics which encodes 'controversial normative assumptions' can be correct. Charlow (2014) does give several empirical arguments for this conclusion which are distinct from the high-level methodological claim that I discuss here. See Lassiter 2014c for a detailed reply to these, and to the empirical criticisms that Cariani (2014) levels at the expected-value semantics specifically. Charlow (2014) follows Carr in rejecting the expected-value theory (along with any other 'contentful' semantics for *ought*) on the grounds that it rules out too much as a matter of meaning. Cariani (2014), too, seems to be sympathetic with the idea that there is a methodological advantage for 'neutral' semantic theories (though he ultimately focuses on empirical questions). All three opt for much less restrictive theories which can generate the predictions of the expected-value semantics among many others. (Semantic theories with similarly weak predictive power, where the heavy lifting is done by relatively unconstrained contextual parameters, have also been proposed recently by others without endorsement of the objection under consideration: see, for example, von Fintel 2012; Dowell 2013.)

In this section I will consider methodological objections of this type, attempting to reconstruct and evaluate an explicit argument in their favor. One motivation for doing so is of course that, if there were an *a priori* advantage for weaker theories of deontic language, this would problematize the semantics that I have just sketched. More broadly, though, I am motivated by what seems to be to be a puzzling fondness for concepts like 'flexibility', 'neutrality', or 'lack of commitments of type X' in some of the recent literature on deontic modality. In other areas of science, flexibility is not be considered to be a theoretical virtue: a theory which is noncommittal about some empirical domain loses the competition with a competing theory which makes specific predictions, unless either (a) the latter can be falsified empirically, or (b) the noncommittal theory gains enough in simplicity to make up for its inferior predictive abilities. We should be using the same criteria for success in evaluating theories of the semantics of deontic modals.

With this in mind, I will now try to construct an explicit *a priori* argument against building 'controversial normative assumptions' into the semantics. I'll argue that the only plausible way to understand this argument is not as a specific competing theory of deontic modals; instead, it is a broad philosophical claim about the nature of lexical semantics which also applies to epistemic modals, scalar adjectives and many other kinds of expressions. In the end, I have no argument against the conclusion of the argument on this reading, but I will argue that it has little practical interest for theorists working to understand the meaning and use of deontic modals.

What is the object that we have been modeling in this paper? Are we interested in devising

- a theory that will optimally capture (for example) *native speakers'* intuitions about entailment and sensicality, where the relevant notion of 'optimality' is sensitive to standard norms of scientific theory-building; or
- a theory which is neutral between all conceivable *theoretical* positions on these topics?

These are obviously not the same thing. In linguistics, there has been a great deal of disagreement about the best grammar formalisms to use in modeling the syntax of natural languages. Some theorists opt for highly restrictive formalisms, while others advocate much more expressive alternatives. The arguments proceed along the usual lines, with a general preference for theories which are as restrictive as possible, but new empirical data sometimes motivates expanded expressive power. No one in this debate would contend that we should use a syntactic formalism which is so powerful that it can express any theory that a reasonable *theoretician* could advocate.

On one reading, the argument about building normative commitments into semantics is analogous to the idea that the best formalism for syntax is one which rules out no *prima facie* plausible syntactic theory. As the argument would go, we should not adopt a semantics which is unable to express the claims of some theory — say, deontology — because there are reasonable, English-speaking theorists who advocate this theory and use 'ought' when discussing ethics with the intention of expressing its *ought*. These theories might be wrong, but surely we do not want them to come out as analytically false!

As a constraint on theory construction, the idea that we should not rule out any plausible theory does not strike me as more reasonable when we are studying the lexical semantics of deontic modals than it is when we are doing syntax. The object that we are modeling is not the range of possible reflective theoretical positions, but the intuitions of native speakers about (for example) entailment, sensicality, truth-value in a specified context. (Such a methodological distinction between syntax and lexical semantics *would* be reasonable if we had reason to think that theorists can reliably introspect the meanings of words in languages that they speak natively. But this is clearly not the case.)

Fixing the intuitions and behavior of (non-theorist) native speakers as the object that we are modeling, then, the conceptual argument reaches another choice point: how broadly applicable is it? Perhaps we might think that deontic modals are special — that, here alone, the *conceivability* of a person who disagrees with some semantic theory indicates that the theory is incorrect. But

I don't know of anyone who has explicitly advocated this idea, and I don't see how an argument for it would go. The alternative is to apply conceivability as a discovery procedure for meaning in general: any piece of information which could conceivably fail to hold is therefore not lexical in nature. This version of the *a priori* argument seems to be essentially equivalent to Moore's Open Question argument. (Charlow (2014) makes this connection for Carr's (2012) argument in particular.)

For example, consider the probabilistic semantics for the epistemic adjective *likely* that was discussed above. Extending the *a priori* argument to epistemic language, we might suppose that the probabilistic theory cannot be correct because it rules out patterns of reasoning that we can imagine an English speaker employing when drawing inferences involving *likely*. (Whether anyone actually *does* violate the predictions of the theory is beside the point for the purposes of this argument.) We would conclude that the semantics should be weakened to allow the likelihood scale to have any conceivable structure: maybe all that is conceptually necessary is transitivity and boundedness, or maybe this is still too much.

Adopting this logic generally would similarly lead us to consign the systematic facts about additive and intermediate adjectives discussed earlier in this paper to 'mere reasoning' — this time, intuitive reasoning about the structure of scalar concepts like heat, weight, danger, and happiness. I suspect that we could use this logic to motivate transferring most or all of what is known about lexical semantics into the realm of concepts or reasoning, on the grounds that there are few beliefs that a person *must* have, and few inferences a person *must* draw, merely by virtue of being a competent speaker of the language. In the limit, we would treat lexical items as mere pointers to aspects of our rich cognitive lives. I hasten to add that this is not an unreasonable position: something like it is adopted by theorists as diverse as Fodor & Lepore (1998), linguists working in the tradition of Distributed Morphology (Halle & Marantz 1993), and Goodman & Lassiter (2014) in their probabilistic theory of language understanding. Crucially, though, this version of the *a priori* argument, if successful, is not a targeted attack on a specific theory of any domain of meaning, but a general relabeling of the enterprise of lexical semantics. The phenomena remain to be explained, and the empirical import of this reclassification is unclear at best.

This discussion connects up directly with some familiar and important themes in philosophy of language. Consider the linguistic and inferential evidence that motivated us to conclude that *heavy* resides on an additive scale. This evidence is intuitively clear to users of English, and so the additivity of weights is a stable piece of implicit knowledge of on the part of English speakers which informs their use and interpretation of the adjectives *heavy* and *light*, the verb *weigh*, the noun *weight*, etc. But what part of their knowledge is this? Is it a piece of linguistic knowledge? Something we know about the concept of heaviness? A stable pattern of reasoning that affects some but not all concepts? A stable component of an intuitive physical theory shared by many (or even all) English speakers?

I have no idea; I also don't know whether these questions make any sense, since they presuppose that there are sharp boundaries between lexical meaning, concepts, reasoning, and the rest of our knowledge and beliefs about the world. What is being asked for here is a clear demarcation between linguistic and non-linguistic information of the type that Quine (1951) famously rejected, arguing that it relies on a fundamental confusion:

> I do not know whether the statement ‘Everything green is extended’ is analytic. Now does my indecision over this example really betray an incomplete understanding, an incomplete grasp of the ‘meanings’, of ‘green’ and ‘extended’? I think not. The trouble is not with ‘green’ or ‘extended’, but with ‘analytic’. (p.31)

Perhaps Quine is incorrect, and there is a meaningful question to be asked about whether decision-theoretic concepts form part of the **meaning** of *good* and *ought* — as opposed to forming a stable part of English speakers’ concept of goodness or of their *good*- and *ought*-involving reasoning patterns. If so, it is sufficiently difficult to draw these boundaries on *a priori* grounds that the fact of their existence cannot be used as a reason to reject any specific theoretical position. But it is also possible that the objection relies on an incorrect conception of the enterprise of lexical semantics, and that this activity forms a continuum with the study of reasoning, concepts, and other aspects of people’s intuitive theories of the world. On this more radical conception, the question of whether these inferential patterns are generated by something that resides in the ‘language box’ is deeply misconceived.

In either case, *a priori* considerations about what should or should not appear in lexical entries are at best indecisive; what will win the day are empirical tests of theories’ predictions, considerations of coherence with our best theories of related domains, and a general preference for the most restrictive theories that are compatible with the empirical facts.See Cariani 2014 for some objections along these lines to the expected-value semantics for *ought*, and Charlow 2014 for some general arguments against ‘contentful’ theories of *ought*. I respond to both sets of arguments in detail in Lassiter 2014c. As far as I can see, the current balance of evidence on these dimensions favors the semantic proposal sketched here.

Daniel Lassiter
Department of Linguistics
Stanford University
450 Serra Mall, building 460
Stanford, CA 94305
danlassiter@stanford.edu

References

Cariani, Fabrizio. 2013. “Ought” and resolution semantics. *Noûs* 47. 534–558.

Cariani, Fabrizio. 2014. Deontic modals and probabilities: One theory to rule them all? In Nate Charlow & Matthew Chrisman (eds.), *Deontic Modals*, Oxford University Press.

Cariani, Fabrizio, Stefan Kaufmann & Magdalena Kaufmann. 2013. Deliberative modality under epistemic uncertainty. *Linguistics and Philosophy* 36. 225–259.

Carlson, Erik. 2014. ‘Good’ in terms of ‘better’. *Noûs* .

Carr, Jennifer. 2012. Deontic modals without decision theory. In *Proceedings of Sinn und Bedeutung 17*, 167–182.

Charlow, Nate. 2014. Decision theory: Yes! Truth conditions: No! In Nate Charlow & Matthew Chrisman (eds.), *Deontic Modals*, Oxford University Press.

Chisholm, Roderick M. & Ernest Sosa. 1966. On the logic of "intrinsically better". *American Philosophical Quarterly* 3(3). 244–249.

Chrisman, Matthew. 2012. 'Ought' and control. *Australasian Journal of Philosophy* 90(3). 433–451.

Cresswell, Max J. 1976. The semantics of degree. In B. Partee (ed.), *Montague Grammar*, 261–292. Academic Press.

Dowell, J. L. 2013. Flexible contextualism about deontic modals: A puzzle about information-sensitivity. *Inquiry* 56(2-3). 149–178.

Finlay, Stephen & Justin Snedegar. 2014. One ought too many. *Philosophy and Phenomenological Research* 89(1). 102–124.

von Fintel, K. 2012. The best we can (expect to) get? Challenges to the classic semantics for deontic modals. Paper presented at the Central meeting of the American Philosophical Association, February 17. http://mit.edu/fintel/fintel-2012-apa-ought.pdf.

Fodor, Jerry A & Ernie Lepore. 1998. The emptiness of the lexicon: Reflections on James Pustejovsky's *The Generative Lexicon*. *Linguistic Inquiry* 29(2). 269–288.

Goble, Lou. 1996. Utilitarian deontic logic. *Philosophical Studies* 82(3). 317–357.

Goodman, Noah D. & Daniel Lassiter. 2014. Probabilistic semantics and pragmatics: Uncertainty in language and thought. In Shalom Lappin & Chris Fox (eds.), *Handbook of Contemporary Semantic Theory*, Wiley-Blackwell 2nd edn.

Gray, John, Andrew McPherson & David Raffe. 2012. *Reconstructions of secondary education: Theory, myth and practice since the second world war*. Routledge.

Gustafsson, Johan E. 2014. Neither 'good' in terms of 'better' nor 'better' in terms of 'good'. *Noûs* 48(3). 466–473. doi:10.1111/nous.12038. http://dx.doi.org/10.1111/nous.12038.

Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Kenneth L. Hale & Samuel J. Keyser (eds.), *The view from building 20*, 111–176. MIT Press.

Halmos, Paul & Steven Givant. 1998. *Logic as algebra* 21. Cambridge University Press.

Hansson, Sven Ove. 1990. Defining "good" and "bad" in terms of "better". *Notre Dame Journal of Formal Logic* 31(1). 136–149.

Hare, R.M. 1967. Some alleged differences between imperatives and indicatives. *Mind* 76(303). 309.

Holliday, Wesley H. & Thomas F. Icard. 2013. Measure semantics and qualitative semantics for epistemic modals. In Todd Snider (ed.), *Semantics and Linguistic Theory (SALT) 23*, Ithaca, NY: CLC Publications.

Huddleston, Rodney & Geoffrey K. Pullum. 2002. *The Cambridge Grammar of English*. Cambridge University Press.

Jackson, Frank. 1985. On the semantics and logic of obligation. *Mind* 94(374). 177–195.

Jackson, Frank. 1991. Decision-theoretic consequentialism and the nearest and dearest objection. *Ethics* 101(3). 461–482.

Jackson, Frank. 2014. Procrastinate revisited *Pacific Philosophical Quarterly* (this issue).

Jackson, Frank & Robert Pargetter. 1986. Oughts, options, and actualism. *The Philosophical Review* 95(2). 233–255.

Jeffrey, Richard C. 1965. *The logic of decision*. University of Chicago Press.

Keynes, John Maynard. 1921. *A Treatise on Probability*. Macmillan.

Kolmogorov, Andrey. 1933. *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Julius Springer.

Kolodny, Niko & John MacFarlane. 2010. Ifs and oughts. *Journal of Philosophy* 107(3). 115–143.

Krantz, David H., R. Duncan Luce, Patrick Suppes & Amos Tversky. 1971. *Foundations of Measurement*. Academic Press.

Kratzer, Angelika. 1977. What 'must' and 'can' must and can mean. *Linguistics and Philosophy* 1(3). 337–355.

Kratzer, Angelika. 1981. The notional category of modality. In Eikmeyer & Rieser (eds.), *Words, Worlds, and Contexts*, 38–74. de Gruyter.

Kratzer, Angelika. 1991. Modality. In von Stechow & Wunderlich (eds.), *Semantics: An international handbook of contemporary research*, de Gruyter.

Kratzer, Angelika. 2012. *Modality and Conditionals: New and Revised Perspectives*. Oxford University Press.

Krifka, Manfred. 1989. Nominal reference, temporal constitution and quantification in event semantics. *Semantics and contextual expression* 75. 115.

Lassiter, Daniel. 2010. Gradable epistemic modals, probability, and scale structure. In Nan Li & David Lutz (eds.), *Semantics & Linguistic Theory (SALT) 20*, 197–215. CLC Publications.

Lassiter, Daniel. 2011. *Measurement and Modality: The Scalar Basis of Modal Semantics*: New York University dissertation.

Lassiter, Daniel. 2014a. Adjectival modification and gradation. In Shalom Lappin & Chris Fox (eds.), *Handbook of Contemporary Semantic Theory*, Wiley-Blackwell 2nd edn.

Lassiter, Daniel. 2014b. Epistemic comparison, models of uncertainty, and the disjunction puzzle. *Journal of Semantics*.

Lassiter, Daniel. 2014c. Linguistic and philosophical considerations on Bayesian semantics. To appear in M. Chrisman & N. Charlow (eds.), *Deontic Modals*. Oxford University Press.

Lassiter, Daniel. 2014d. *Measurement and Modality: The Scalar Basis of Modal Semantics*. Oxford University Press. To appear.

Lewis, David. 1973. *Counterfactuals*. Harvard University Press.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In R. Bäuerle, C. Schwarze & A. von Stechow (eds.), *Meaning, use and interpretation of language*, vol. 21, 302–323. Walter de Gruyter.

Nerbonne, John. 1995. Nominal comparatives and generalized quantifiers. *Journal of Logic, Language and Information* 4(4). 273–300.

Quine, Willard van Orman. 1951. Two dogmas of empiricism. *The Philosophical Review* 60(1).

Regan, Donald. 1980. *Utilitarianism and Co-operation*. Oxford University Press.

van Rooij, Robert. 2011. Measurement, and interadjective comparisons. *Journal of Semantics* 28(3). 335–358.

Sassoon, Galit. 2010. Measurement theory in linguistics. *Synthese* 174(1). 151–180.

Scha, Remko & Yoad Winter. 2014. Plurals. In Shalom Lappin & Chris Fox (eds.), *Handbook of Contemporary Semantic Theory*, Wiley-Blackwell 2nd edn.

Schroeder, Mark. 2011. Ought, agents, and actions. *Philosophical Review* 120(1). 1–41.

Swanson, Eric. 2006. *Interactions With Context*: MIT dissertation. http://mit.dspace.org/bitstream/handle/1721.1/37356/123190034.pdf?sequence=1.

Wedgwood, Ralph. 2006. The Meaning of "Ought". *Oxford studies in Metaethics* 1. 127–60.
Yalcin, Seth. 2005. A puzzle about epistemic modals. *MIT Working Papers in Linguistics* 51. 231–272.
Yalcin, Seth. 2007. Epistemic modals. *Mind* 116(464). 983–1026. doi:10.1093/mind/fzm983.
Yalcin, Seth. 2010. Probability Operators. *Philosophy Compass* 5(11). 916–937.