Attitude Reports
Thomas Grano (tgrano@indiana.edu)
For: "Key Topics in Semantics and Pragmatics", Cambridge University Press
Ms. date: September 20, 2017

Provisional chapter outline:

1. Introduction
2. Attitude reports in possible worlds semantics
3. Frege's puzzle
4. Attitudes *de re*
5. Attitudes *de se*
6. Beyond belief: Desire reports and the typology of attitude predicates
7. Intensional transitive verbs
8. Negative raising

Contents

6 Beyond belief: Desire reports and the typology of attitude predicates

6.1 Introduction

If we treat *believe*-sentences as instantiating universal quantification over the attitude holder's doxastic alternatives, as in (1), then a reasonable null hypothesis would be that *want*-sentences are just the same except that the quantification is over the attitude holder's bouletic alternatives (i.e., all those worlds in which the attitude holder's desires are satisfied), as in (2).

(1) $[\![\text{Beatrix believes it's raining}]\!]^w = \forall w' \in \text{DOX}_{b,w}$: rain($w'$)
where $\text{DOX}_{x,w} = \{w' \mid w'$ is compatible with x's beliefs in $w\}$

(2) $[\![\text{Beatrix wants it to be raining}]\!]^w = \forall w' \in \text{BOUL}_{b,w}$: rain($w'$)
where $\text{BOUL}_{x,w} = \{w' \mid w'$ is compatible with x's desires in $w\}$

As it turns out, a wealth of evidence points toward the conclusion that (2) is wrong: there are simply too many differences between *believe*-sentences and *want*-sentences, and too many empirical difficulties for the approach in (2), to lend any credence to

such a treatment. What exactly should replace (2) is still a matter of some debate, though we do know quite a bit about the various issues that any adequate proposal must be able to answer to. In what follows, we review some of the more salient issues and the proposals that they have prompted. We then scale up to take a broader look at semantic variation among attitude predicates.

6.2 Teaching Tuesdays and Thursdays: Belief-relativity

We begin with an observation due to Heim (1992) that exposes a problem for treating *want*-sentences as instantiating universal quantification over bouletic alternatives. Heim asks us to consider the sentence in (3).

(3) I want to teach Tuesdays and Thursdays next semester. (Heim 1992:195)

And here is what Heim has to say about this sentence:

> Suppose this sentence is intuitively true as spoken by me today. Is it therefore the case . . . that I teach Tuesdays and Thursdays next semester in all the worlds that are compatible with everything I desire? No. In worlds that are compatible with everything I desire I actually don't teach at all. (Heim 1992:195)

In other words, universal quantification over bouletic alternatives seems to endow *want*-sentences with a semantics that is too strong: such an analysis predicts that we should only be able to use *want*-sentences truthfully when they describe a state of affairs that would hold in those worlds where all of the attitude holder's desires (no matter how wild or farfetched) are fulfilled. And this is simply not the case, as Heim's example attests to. So then the question is: what is the right way of weakening the semantics of *want*-sentences?

Predominant approaches in the formal semantics literature take as their starting point the intuition that the semantics of *want*-sentences is intimately bound up with the attitude holder's beliefs. In Heim's (3), for example, what is intuitively needed is for the semantics to allow us to ignore those worlds incompatible with the attitude holder's beliefs: just considering alternatives that are realistic from the point of the view of the attitude holder (which will restrict us to worlds where the attitude holder does indeed teach next semester), a Tuesday-Thursday teaching schedule is preferable to any other realistically possible teaching schedule.

There are a number of proposals in the literature for cashing out this basic idea; here I will introduce two that have been quite influential, and we will consider refinements and alternatives as we proceed. The first is what I will call the BETTER-WORLDS approach and it is the one that Heim (1992) proposes. The second is what

I will call the BEST-WORLDS approach and it has one of its clearest articulations in von Fintel (1999) (but prior to this can also be found to one degree or another in Kratzer 1981; Giorgi and Pianesi 1997; Portner 1997). In what follows I introduce each in turn.

6.2.1 The better-worlds approach

Heim's better-worlds approach takes inspiration from Stalnaker (1984), who wrote, in a now frequently quoted passage, "wanting something is preferring it to certain relevant alternatives, the relevant alternatives being those possibilities that the agent believes will be realized if he does not get what he want" (p. 89). Taking seriously the *if*-clause in Stalnaker's characterization of what wanting is, Heim proposes that *want*-sentences have a conditional semantics, whereby (4-a) has truth conditions paraphrasable as (4-b).

(4) a. Beatrix wants it to be raining.
b. Beatrix believes that if it is raining, she will be in a more desirable world than if it is not raining.

In order to cash out (4-b) in formal terms, we need a semantics for *believe*-sentences, a semantics for conditionals (*if*-clauses), and a semantics for comparative desirability (i.e., for . . . *be in a more desirable world than*. . .).

For *believe*-sentences, Heim adopts the standard Hintikkan approach of universal quantification over doxastic alternatives. Since this has already been extensively reviewed in previous chapters, we will not dwell on it any more here.

For conditionals, Heim follows an approach combining components of Stalnaker 1968 and Lewis 1973 that works in such a way that a conditional sentence like (5) has the semantics in (6). According to (6), an *if*-clause p is mapped onto a function Sim_w that returns that subset of p-worlds maximally similar to the evaluation world w. If all of these worlds are worlds in which the consequent of the conditional is true, then the conditional sentence is true.

(5) If it is raining, Beatrix will leave.

(6) $\text{Sim}_w(\llbracket\text{it is raining}\rrbracket) \subseteq \llbracket\text{Beatrix will leave}\rrbracket$
where $\text{Sim}_w(p) = \{w'\colon\ w' \in p$ and w' resembles w no less than any other world in $p\}$ (Heim 1992:195)
'Those worlds in which it is raining that are maximally similar to w are all worlds in which Beatrix will leave.'

Finally, for comparative desirability, Heim defines a relation $>_{a,w}$ that compares desirability of worlds relative to an attitude holder a and an evaluation world

w (7-a), and Heim also defines a derivative variant that compares desirability of propositions (sets of worlds) relative to an attitude holder and an evaluation world, as in (7-b).[1]

(7) For any w, w', w'':

a. $w' >_{a,w} w''$ iff w' is more desirable to a in w than w''

b. $X >_{a,w} Y$ iff $w' >_{a,w} w''$ for all $w' \in X$, $w'' \in Y$ (Heim 1992:197)

Putting together the assumed semantics for belief, conditionals, and comparative desirability, (8) is what we end up with for the semantics of the *want*-sentence under discussion. In procedural terms, it says: consider all those worlds compatible with Beatrix's beliefs. At each one of those worlds, check to see whether maximally similar worlds in which it is raining are all more desirable to Beatrix than maximally similar worlds in which it is not raining. If this is the case for all of the belief worlds, then the sentence is true; otherwise, the sentence is false.

(8) $[\![\text{Beatrix wants it to be raining}]\!]^w =$
$\forall w' \in \text{DOX}_{b,w}$: $\text{Sim}_{w'}(\text{raining}) >_{b,w} \text{Sim}_{w'}(\neg\text{raining})$

As it happens, (8) is still not quite Heim's final semantics for *want*-sentences. The reason is that in (8), nothing prevents the Sim function from picking out worlds incompatible with Beatrix's beliefs, because it could be that at least some rain-worlds maximally similar to at least some of Beatrix's belief worlds are not themselves among Beatrix's belief worlds (and likewise for non-rain worlds). A potential reason for wanting to disallow this comes from the minimal trio in (9).

(9) a. Beatrix wants it to be raining.

b. Beatrix wishes that it were raining.

c. Beatrix is glad that it is raining.

What is the difference between *want*, *wish*, and *be glad*? One salient difference is that *wish*-sentences are used to say something about the relative desirability of worlds that are not among the attitude holder's doxastic alternatives, whereas *be glad*-sentences are used to say something about the relative desirability of worlds that are among the attitude holder's doxastic alternatives. Heim proposes that *want*-sentences are somewhere in between, used to report the relative desirability of worlds that partially overlap with the attitude holder's doxastic alternatives; i.e., propositions that the attitude holder considers to be possible (unlike *wish*) but not

[1] Heim (1992) uses the 'less-than' symbol '$<$' for comparative desirability. Here I follow Villalta (2008) in replacing it with the greater-than symbol '$>$', because it is more intuitive for me to think of 'more desirable' as being 'greater than' rather than 'less than'.

guaranteed (unlike *be glad*). Strengthening the analytical connection between desire reports and conditional sentences, Heim notes a striking parallel between the varieties of desire reports exemplified by (9) and the varieties of conditionals exemplified by (10): the indicative conditional in (10-a) is 'want'-like in that rain is considered possible but not guaranteed, the counterfactual conditional in (10-b) is 'wish'-like in that rain is considered to be counterfactual, and the *because*-sentence in (10-c) is 'be glad'-like in that rain is taken for granted.

(10) a. If it is raining, Beatrix is staying inside.
b. If it were raining, Beatrix would be staying inside.
c. Because it is raining, Beatrix is staying inside.

To implement this idea, then, Heim modifies the semantics so that before the proposition named by the complement to *want* and its negative counterpart are run through the Sim function, it is intersected with the attitude holder's doxastic alternatives, as in (11). Since the Sim function can only return subsets of the propositions it applies to, this intersection ensures that the output of the Sim function will only contain worlds compatible with the attitude holder's beliefs.[2]

(11) $[\![\text{Beatrix wants it to be raining}]\!]^w =$
$\forall w' \in \text{DOX}_{b,w}$: $\text{Sim}_{w'}(\text{DOX}_{b,w} \cap \text{raining}) >_{b,w} \text{Sim}_{w'}(\text{DOX}_{b,w} \cap \neg\text{raining})$

Does Heim's semantics account for the Tuesday/Thursday example repeated in (12)? It does: on Heim's approach, (12) asserts that at each of the attitude holder's belief worlds, maximally similar worlds in which she teaches Tuesdays and Thursdays are all more desirable than maximally similar worlds in which she does not teach Tuesdays and Thursdays.

(12) I want to teach Tuesdays and Thursdays next semester.

In Heim's scenario where the attitude holder believes that she will teach no matter what, maximally similar worlds in which she does not teach Tuesdays and Thursdays are worlds in which she teaches on some other schedule, not worlds in which she does not teach at all. So the sentence ends up not saying anything about the absolute desirability of teaching Tuesdays and Thursdays, but rather only about the relative desirability of teaching Tuesdays and Thursdays as opposed to other days

[2]Heim's semantics actually still looks somewhat different from (11), because Heim employs a dynamic semantic framework rather than a static one. One of the reasons Heim uses a dynamic semantics is to help carry out an overarching aim of her paper, which is to account for some facts related to presupposition projection (see section 6.9 below). Since the static/dynamic distinction is orthogonal to our immediate concerns here, I follow Villalta's (2008) translation of Heim's approach into a static framework.

of the week. And this is exactly what we want.

6.2.2 The best-world approach

Having sketched Heim's better-world semantics for *want*-sentences, let's now consider von Fintel's version of the best-worlds approach. Whereas Heim's approach borrows technology from the semantics of conditionals, von Fintel's approach borrows technology from Kratzer's (1981) highly influential doubly relative semantics for modals. By way of background, consider the deontic modal sentence in (13).

(13) Maggie has to pay a fine.

An extremely naive first approximation of a possible worlds analysis of (13) would amount to something like, 'All those worlds in which the rules are obeyed are worlds in which Maggie pays a fine.' The problem with this is that typically one pays a fine because one has broken a rule. So in all of those worlds in which the rules are obeyed, it is not in fact the case that Maggie pays a fine, because in those worlds, Maggie's fine-prompting infraction never occurred. Kratzer's (1981) solution to this problem was for modals to be interpreted relative to two parameters: a modal base and an ordering source. Then in calculating the meaning of (13), we proceed something like this: consider all of those worlds in which the circumstances of this world up until now (including Maggie's fine-prompting infraction) have occurred. This is the modal base. Then, order those worlds according to how closely they approximate the deontically "ideal" worlds in which all rules are obeyed. This is the ordering source. Those worlds that most closely approximate this ideal set are all worlds in which Maggie pays a fine.

Technically, this is achieved by defining an ordering relation as in (14). Read $<_P$ as 'better than, relative to the set of ideals named by P'. Then, according to (14), a world w is better than a world w' relative to P iff the ideals that are realized in w' is a proper subset of the ideals that are realized in w; or in other words, iff w realizes more ideals than w'. Then we define a function $\max_P$ that takes a set of worlds and returns that subset of worlds that come closest to satisfying the ideals named by P, as in (15).

(14) Given a set of worlds X and a set of propositions P: $\forall w, w' \in X$:
$w <_P w'$ iff $\{p \in P\text{: } p(w') = 1\} \subset \{p \in P\text{: } p(w) = 1\}$ (von Fintel 1999:115)

(15) Given a set of worlds X and a strict partial order $<_P$: $\forall X \subseteq W$:
$\max_P(X) = \{w \in X\text{: } \neg\exists w' \in X\text{: } w' <_P w\}$ (von Fintel 1999:116)

Applying this idea to *want*-sentences, the proposal is that the modal base is the set of worlds compatible with the attitude holder's beliefs (doxastic alterna-

tives), and the ordering source is given by the attitude holder's desires (bouletic alternatives). This is illustrated in (16). In procedural terms, it tells us: look at all those worlds compatible with Beatrix's beliefs. Then, pick out just that subset of her belief worlds that best satisfy her desires. If, in all of those best worlds, it is raining, then the sentence is true; if not, the sentence is false.

(16) $[\![\text{Beatrix wants it to be raining}]\!]^w = \forall w'\ \max_{BOUL_{b,w}}(\text{DOX}_{b,w})$: raining($w'$)

Like Heim's better-worlds approach, the best-worlds approach also accounts for the Tuesday/Thursday example: we look at all of the worlds compatible with the attitude holder's beliefs, and then consider that subset that best satisfies the attitude holder's desires. If the attitude holder teaches Tuesdays and Thursdays in all of those worlds, then the sentence is true, and this is fully consistent with a scenario where there are even better worlds, crucially not among the attitude holder's belief worlds, in which the attitude holder does not teach at all.

Having now sketched both the better-worlds and the best-worlds approach to *want*-sentences, we turn next to some further empirical considerations, some of which will prompt some refinements and some of which may bear on which (if either) of the two approaches is correct.

6.3 Believing without wanting: A doxastic presupposition for want-*sentences*

An immediate problem for both the better-worlds approach and the best-worlds approach to *want*-sentences as sketched above is the prediction that a *believes* p entails a *wants* p (for arbitrary values of a and p). This is a bad result: after all, if I *believe* that I will die in ten days, this does not by any stretch of the imagination entail that I *want* to die in ten days.

Consider first the better-worlds approach. Suppose that Beatrix believes that it is raining. This fact alone is sufficient to verify (17), because $\text{DOX}_{b,w} \cap \neg$rain and therefore $\text{Sim}_{w'}(\text{DOX}_{b,w} \cap \neg\text{rain})$ denotes the empty set, and it is trivially the case that every world in $\text{Sim}_{w'}(\text{DOX}_{b,w} \cap \text{rain})$ is more desirable to Beatrix than any world in the empty set, because there are no worlds in the empty set. In fact, by parity of reasoning, the better-worlds approach also erroneously predicts that a *believes* $\neg p$ entails a *wants* p: suppose Beatrix believes that it is not raining. Then, $\text{Sim}_{w'}(\text{DOX}_{b,w} \cap \text{rain})$ denotes the empty set, and it is trivially the case that all of the worlds in the empty set are more desirable to Beatrix than any of the worlds in $\text{Sim}_{w'}(\text{DOX}_{b,w} \cap \neg\text{rain})$.

(17) $[\![\text{Beatrix wants it to be raining}]\!]^w =$
$\forall w' \in \text{DOX}_{b,w}$: $\text{Sim}_{w'}(\text{DOX}_{b,w} \cap \text{rain}) >_{b,w} \text{Sim}_{w'}\ (\text{DOX}_{b,w} \cap \neg\text{rain})$

Heim's solution to this problem is a rather gentle modification: we tack a condition onto the Sim_w function, so that when it applies to a set of possible worlds p, the output is defined only if p is not the empty set:

(18) p is in the domain of Sim_w only if $p \neq \emptyset$; where defined, $\text{Sim}_w(p)$ = $\{w'$: $w' \in p$ and w' resembles w no less than any other world in $p\}$ (Heim 1992:198)

This ends up having as a consequence that if Beatrix believes that it is raining, or if Beatrix believes that it is not raining, then the sentence *Beatrix wants it to be raining* is neither true nor false but simply undefined. More generally, the consequence is that a *wants* p presupposes that p is consistent with but not entailed by a's beliefs. (Though see section 6.4 for some further complications.) So the erroneous prediction about entailment relations is blocked by treating the entailed *want*-sentences as presupposition failures.

Consider now the best-worlds approach. Suppose again that Beatrix believes that it is raining. Just like on the better-worlds approach, this belief alone is sufficient to verify (19): (19) tell us that the *want*-sentence is true iff some subset of Beatrix's belief worlds (namely those that best satisfy her desires) are worlds in which it is raining. But if Beatrix believes that it is raining (i.e., if *all* of her belief worlds are such that it is raining), then it is trivially the case that any subset of those belief worlds are also all such that it is raining: the restriction of a universal quantifier is a downward entailing environment.

(19) $[\![\text{Beatrix wants it to be raining}]\!]^w = \forall w'\ \max_{BOUL_{b,w}}(\text{DOX}_{b,w})$: rain($w'$)

von Fintel's (1999) solution is essentially to translate Heim's solution into the best-worlds framework: we do this by tacking a condition onto *want* itself, such that after all its arguments are fed to it, its output is defined only if $\text{DOX}_{a,w} \cap p \neq \emptyset$ (a considers p possible) and $\text{DOX}_{a,w} - p \neq \emptyset$ (a does not consider p certain):

(20) $[\![\text{Beatrix wants it to be raining}]\!]^w$ is defined only if $\text{DOX}_{b,w} \cap \text{rain} \neq \emptyset$ and $\text{DOX}_{b,w} - \text{rain} \neq \emptyset$.
If defined, $[\![\text{Beatrix wants it to be raining}]\!]^w$ = $\forall w'\ \max_{BOUL_{(b,w)}}(\text{DOX}_{b,w})$: rain($w'$)

This correction has the same consequence as it does on the better-worlds approach.

In order for Heim's and von Fintel's solutions to be theoretically satisfying, it ought to be the case that the proposed presuppositions have broader applicability than just *want*-sentences. And both solutions seem to succeed by this metric: Heim's solution predicts that (indicative) conditional sentences in general should presuppose that the antecedent of the conditional is possible but not certain, which

seems right: (21) seems to presuppose that rain is a possibility (otherwise it would have been more appropriate to use a counterfactual conditional) but not guaranteed (otherwise it would have been more appropriate to use *because* rather than *if*). And von Fintel's solution predicts that modal statements in general should presuppose that the modal base worlds are consistent with but do not entail the prejacent. This seems right as well: (22) seems to presuppose that in some but not all worlds compatible with the circumstances up until the reference time, John pays a fine. In the modal semantics literature this is sometimes known as the Diversity Condition (see e.g. Condoravdi 2002).

(21) If it's raining, then I won't need to water the plants tomorrow.

(22) John had to pay a fine.

6.4 Going to the movies and never-ending weekends: Refining the presupposition

Let's take stock. We started out with some considerations pointing to the conclusion that *want*-sentences incorporate belief semantics in an intimate way, and we looked at two implementations of this idea. These implementations made the desire/belief connection a little *too* intimate, erroneously predicting that beliefs trivially entail corresponding desires. To remedy this, we appealed to presupposition: *want*-sentences presuppose that the attitude holder considers the desired proposition possible but not guaranteed. This presupposition offsets the belief/desire intimacy, imposing a healthy distance between the two, in a way that seems to get the balance just right. Or does it?

Heim herself noticed that both the 'possible' and the 'not guaranteed' components of the proposed presupposition seem to face counterexamples. In particular she considered the sentences in (23)–(24). (23) is problematic insofar as it could be used felicitously in a situation where John believes he will go to the movies tonight, in apparent violation of the 'not guaranteed' presupposition. And (24) is problematic insofar as it is felicitous even in the presence of an explicit denial of possibility, in apparent violation of the 'possible' presupposition.

(23) (John hired a babysitter because) he wants to go to the movies tonight. (Heim 1992:199)

(24) I want this weekend to last forever. (But I know, of course, that it will be over in a few hours.) (Heim 1992:199)

In response to the problem raised by (23), Heim proposes a core revision to her semantics: rather than having *want*-sentences incorporate DOX, as in (25), Heim suggests, we should analyze *want*-sentences as incorporating a superset of

DOX (what von Fintel, in later adopting the revision, dubs DOX*), defined by those worlds that the attitude holder believes are possible irrespective of how the attitude holder may choose to act, as in (26). Heim observes that we routinely use *want*-sentences to report on intentions to act, and intended outcomes are typically not just believed to be possible but rather believed to be certain, a certainty grounded by the intention itself: if I intend to go out tonight, then (barring bizarre circumstances), of course I believe I will, because the intention is a commitment to see to it that I do. (This is well studied in the philosophy literature: see especially Bratman 1987, and see also Grano 2017b on the semantics of intention reports.)

(25) $\text{DOX}_{x,w} = \{w' \mid w' \text{ is compatible with } x\text{'s beliefs in } w\}$

(26) $\text{DOX*}_{x,w} = \{w' \mid w' \text{ is compatible with everything } x \text{ believes to be the case in } w \text{ no matter how he chooses to act}\}$

This solution thus pries apart the belief/desire connection every so slightly more. In (23), John going to the movies tonight is presumably entailed by $\text{DOX}_{j,w}$, but crucially not by $\text{DOX*}_{j,w}$, given that John could choose to act otherwise. The presupposition failure is thereby defused. Whether we should be happy with this solution depends in part on what kinds of broader predictions it makes and whether it is needed elsewhere in the grammar, a question that as far as I know has not been extensively investigated.

And what about the problem of outcomes not believed to be possible? Here Heim is noncommittal: she suggests that either they too involve some superset of $\text{DOX}_{x,w}$ too weak to entail impossibility, or that on some level the attitude holder really does consider the outcome possible: as she puts it, "The reasonable part of me knows and is resigned to the fact that time passes, but the primitive creature of passion has lots sight of it" (p. 200). As far as I know, subsequent scholarship has not followed up on this problem very extensively. One partial exception is Rubinstein (2012), who suggests that "possibilities that are relevant for a desire statement may be possibilities that are circumstantially accessible, yet doxastically inaccessible" (p. 116). But the suggestion is not developed enough to be fully assessed and it is not entirely clear what would make a never-ending weekend circumstantially accessible. It has also been observed in the literature that *hope* really does seem to carry the presupposition of possibility that Heim proposed for *want*, as evidenced by examples such as (27), and *hope* is indeed odd in the never-ending weekend sentence (28) (see also section 6.10 below). More research is needed here.

(27) I want/??hope to build a perpetual motion machine. (Portner and Rubinstein 2012:472)

(28) ??I hope this weekend lasts forever. (But I know, of course, that it will be

over in a few hours.)

6.5 Getting well, teaching, and riding the Concorde: Monotonicity

One important way in which the better-worlds and the best-worlds approaches to *want*-sentences pull apart from each other empirically is in their predictions about monotonicity. As stressed by von Fintel (1999), the best-worlds approach makes *want*-sentences monotonic; more specifically, it predicts that if a *wants* p is true, then for any proposition q such that p entails q, a *wants* q is also true. (I.e., a *wants* p is upward-entailing on p.) The reason for this is simple: on the best-worlds approach, p is mapped onto the scope of a universal quantifier (the restriction being the attitude holder's most desirable doxastic alternatives). So upward-entailingness is guaranteed for the same reason that *All dogs are mammals* and *All mammals are animals* together entail *All dogs are animals*: the subset relation involved in universal quantification is a transitive one. The better-worlds approach, on the other hand, does not predict upward-entailingness: just because, at each doxastic alternative, all the most similar p-worlds are better than all the most similar $\neg p$-worlds, there is no guarantee that any given superset of p-worlds (call them q-worlds) are such that, at each doxastic alternative, the most similar q-worlds all better than all the most similar $\neg q$-worlds. To venture an analogy from the domain of individuals (individuals are generally far easier to reason about than possible worlds), suppose (counterfactually) that all semanticists are smarter than all non-semanticists, and (perhaps truthfully) that all semanticists are linguists. It does not then follow that all linguists are smarter than all non-linguists. So the predictions of the two approaches are clear. Which is right? As it turns out, deciding whether *want*-sentences are in fact upward-entailing or not is a rather delicate matter.

Let's start with an easy set of cases that both approaches are able to handle. In a passage quoted by both Heim (1992) and von Fintel (1999), Stalnaker (1984) writes as follows:

> Suppose I am sick. I want to get well. But getting well entails having been sick, and I do not want to have been sick. Suppose there was a murder. I want to know who committed the murder. But my knowing who committed the murder entails that the murder was committed, and I never wanted the murder to have been committed. (Stalnaker 1984:89)

These cases look like failures of upward entailingness and hence would seem to support the better-worlds approach over the best-worlds approach. But note a crucial feature of both examples: in the first, the attitude holder *believes* that he is sick and

in the second, the attitude holder *believes* a murder was committed. Hence a proponent of the best-worlds approach would be justified in saying that what's wrong with the inferences is not that they are false but rather that they incur a presupposition failure: as seen in section 6.3, we have independent reasons for thinking that *want*-sentences presuppose that the attitude holder is not convinced that the desired outcome is guaranteed. So examples like Stalnaker's are not helpful in deciding between the best-worlds and the better-worlds approaches.

Now let's turn to another kind of example, discussed by Heim (1992), building on a similar example by Asher (1987). Consider the following pair of sentences:

(29) a. Nicholas wants a free trip on the Concorde.
b. Nicholas wants a trip on the Concorde. (Heim 1992:194, based on Asher 1987:171)

Heim writes of these sentences:

> [I]magine that Nicholas is not willing to pay the $3,000 that he believes it would cost him if he flew to Paris on the Concorde, but he would love to fly on the Concorde if he could get the trip for free. Under these circumstances [(29-a)] is true, yet [(29-b)] is false, despite the fact that taking a free trip on the Concorde, of course, implies taking a trip on the Concorde. (Heim 1992:194)

Is this a knock-down argument against the best-worlds approach? It turns out not to be, because a best-worlds proponent could reply (as von Fintel does) that in testing for upward-entailingness one must control for changes to the context. If Nicholas believes that flying the Concorde costs $3,000, then in reasoning about what he wants, we consider only those worlds where this belief is true, and because of this none of the most desirable worlds are ones where he takes a trip on the Concorde. But if the context then changes so that a free trip on the Concorde is possible, this expands the set of worlds compatible with Nicholas's beliefs and therefore affects the computation of what he wants: in this new expanded set, the most desirable worlds are indeed ones where he takes a trip on the Concorde, namely some subset of those where he does so for free.

In fact, von Fintel argues not only that the Concorde example is consistent with the best-worlds approach but even more strongly that only the best-worlds approach and not the better-worlds approach accurately predicts that paired desire reports like (30) or (31) in fact sound contradictory. In other words, if we do not introduce any contextual factors that affect the set of worlds being quantified over from one desire report to the next, then *want*-sentences seem to be upward entailing after all.

(30) #Nicholas does not want a trip on the Concorde but he does want a free trip on the Concorde.

(31) #Nicholas wants a free trip on the Concorde but he does not want a trip on the Concorde.

Other authors have expressed similar intuitions pointing toward the correctness of an upward-entailing analysis of *want*-sentences. Crnič (2011:172), for example, cites (32-a–b) as valid inferences, and Condoravdi and Lauer (2016:32) contend that (33) "seems to be a valid piece of reasoning" (though they ultimately propose a non-monotonic semantics for *want* and suggest that (33) involves allowing indefinites to take wide scope and quantify over properties; see also section 6.6 below).[3]

(32) a. I want to read five books → I want to read one book
b. I want to meet Mary and Sue → I want to meet Mary

(33) John wants to get a poodle, so he wants to get a dog.

In a similar vein, Fara (2013), reacting to a view popular in philosophy that the content of a desire report specifies exactly what it would take for the desire to be satisfied, discusses examples like (34)–(35). As Fara points out, (34) could intuitively be truthful in a situation where "Fiona wants to catch a fish that's big enough to make a meal; a minnow will not do" (p. 250); similarly, (35) could intuitively be truthful in a situation where "Charlotte wants enough champagne to feel it go to her head; a thimbleful will not do" (p. 250).

(34) Fiona wants to catch a fish.

(35) Charlotte wants to have some champagne.

Although Fara does not explicitly connect these observations to the debate over upward-entailingness, they are directly relevant to it: if *want*-sentences are upward-entailing as the best-worlds approach would have it, then their content merely spec-

[3]Crnič (2011) also discusses one systematic source of apparent counterexamples to upward entailingness, namely disjunction introduction. For example, (i-a) seems not to entail (i-b), against the expectations of an upward entailing analysis. This observation has a long history in the deontic logic literature, where it is known as Ross's Paradox, after Ross 1941. Crnič suggests a way of reconciling this with an upward-entailing analysis, but assessing it here would take us too far afield into the semantics of free choice and exhaustification, so I encourage interested readers to consult Crnič 2011.

(i) a. John wants to send this letter.
b. John wants to send this letter or burn it.

ifies NECESSARY CONDITIONS for what it would take to satisfy the desire, rather than necessary and sufficient conditions. So Fara's observations seem to support an upward-entailing analysis.

That being said, Fara's reasoning has not gone unchallenged: in a reply paper, Braun (2015) suggests that what could be going on in examples like (34) is that Fiona wants to catch a meal-sized fish, and then reflects on this desire and realizes that in order to catch a meal-sized fish, she has to catch a fish, and thereby forms another desire, namely the one stated in (34). If she were to then catch a minnow, the desire named in (34) would in fact be satisfied, but her desire to catch a meal-sized fish would not be. In other words, we need to be careful not to make entailments out of what could instead be assumptions about how rational agents (should) reason. And this caution should of course look familiar from studying belief reports.

6.6 Visiting Paris and Rome: Conjunction introduction and conflicting desires

Recall from Chapter 2 two basic logical properties of *believe*-sentences on the Hintikkan approach: first, *believe*-sentences obey conjunction introduction, in that, for an arbitrary attitude holder a and arbitrary propositions p and q, a *believes* p and a *believes* q together imply a *believes* $[p \wedge q]$. And second, paired *believe*-sentences whose contents contradict each other are themselves contradictory. As we discussed there, these are arguably desirable features of a semantics for belief reports, at least insofar as we want the semantics to be normative on how we think rational agents ought to behave.

But the situation is different for *want*-sentences. Levinson (2003) asks us to imagine an individual John who is planning his summer vacation. He considers Paris and Rome to both be desirable travel destinations, but does not want to expend the time or money needed to visit both cities, even though he believes that visiting both cities is in principle possible. In such a scenario, (36) seems truthful and felicitous, apparently showing us that *want*-sentences do not obey conjunction introduction. And crucially, (36) does not even portray John as irrational. Thus even if we wanted our attitude semantics to be normative on rationality, we would still not want *want*-sentences to obey conjunction introduction.

(36) John wants to visit Paris and John wants to visit Rome, but John does not want to visit Paris and Rome.

Consider now, as Levinson (2003) does, a slight variant on the previous scenario. As before, John considers Paris and Rome to both be desirable travel destinations, but this time he outright lacks the monetary resources to visit both

cities and consequently believes that visiting both cities would be impossible. In this scenario as well, it is possible to say, without contradiction, that John wants to visit Paris and John wants to visit Rome. In other words, the semantics of *want*-sentences needs to allow for conflicting desires without resulting in contradiction. To reinforce this point, here are some other examples of conflicting desires culled from the literature:

(37) Robert wants to marry Jane, but Robert also wants to marry Sue (and bigamy is not possible). (adapted from Portner 2005:164)

(38) I want it to rain tomorrow so the picnic gets canceled, but I (also) want it to be sunny tomorrow so I can go hiking. (Condoravdi and Lauer 2016:24)

(39) John wants to move in with his girlfriend, but he also wants to keep living alone. He can't make up his mind. (Condoravdi and Lauer 2016:24)

(40) [I believe that I will teach Tuesdays and Thursdays next semester if and only if I work hard now:]
I want to teach Tuesdays and Thursdays next semester but I do not want to work hard now. (adapted from Villalta 2008:478)

(41) [Al knows that he'll see the concert only if he takes the long drive:]
Al wants to see the concert, but Al doesn't want to take the long drive. (adapted from Phillips-Brown to appear)

The best-worlds approach to *want*-sentences erroneously predicts both that *want*-sentences should obey conjunction introduction and that paired *want*-sentences with contradictory contents should themselves be contradictory. On the best-worlds approach, if John wants to visit Paris, this means that all of his best doxastic alternatives are ones where he visits Paris. And if John wants to visit Rome, all of his best doxastic alternatives are ones where he visits Rome. The only way for this to be the case is if all of his best doxastic alternatives are ones in which he visits both Paris and Rome. This follows from the same reasoning that guarantees that if all dogs are animals and all dogs are quadrupeds, then all dogs are quadruped animals. And by the same token, if there are no doxastic alternatives where John visits both Paris and Rome, then we end up with a logical contradiction.

Does the better-worlds approach fare any better with respect to conjunction introduction and conflicting desires? Levinson (2003:229–230) and Crnič (2011:171) both claim that it does, but it is not so clear to me that this is right. Consider a toy model where John has just four doxastic alternatives:

(42) a. w_1: John visits Paris but does not visit Rome.
b. w_2: John visits Rome but does not visit Paris.

c. w_3: John visits neither Paris nor Rome.
d. w_4: John visits both Paris and Rome.

Given (42), the better-worlds approach tells us that the following holds:

(43) $[\![\text{John wants to visit Paris}]\!]^w$ =
$\forall w' \in \text{DOX}_{j,w}$: $\text{Sim}_{w'}(\text{DOX}_{j,w} \cap \{w_1, w_4\}) >_{j,w} \text{Sim}_{w'}(\text{DOX}_{j,w} \cap \{w_2, w_3\})$

This is in turn equivalent to:

(44) a. $\text{Sim}_{w_1}(\{w_1, w_4\}) >_{j,w} \text{Sim}_{w_1}(\{w_2, w_3\}) \wedge$
b. $\text{Sim}_{w_2}(\{w_1, w_4\}) >_{j,w} \text{Sim}_{w_2}(\{w_2, w_3\}) \wedge$
c. $\text{Sim}_{w_3}(\{w_1, w_4\}) >_{j,w} \text{Sim}_{w_3}(\{w_2, w_3\}) \wedge$
d. $\text{Sim}_{w_4}(\{w_1, w_4\}) >_{j,w} \text{Sim}_{w_4}(\{w_2, w_3\})$

Carrying out the consequences of the Sim functions yields (45), given plausible assumptions about world similarity (worlds are always more similar to themselves than to any other world, worlds where John visits neither Paris nor Rome are more similar to worlds where John visits Paris but not Rome than are worlds where John visits Rome but not Paris, etc.).

(45) a. $w_1 >_{j,w} w_3 \wedge$
b. $w_4 >_{j,w} w_2 \wedge$
c. $w_1 >_{j,w} w_3 \wedge$
d. $w_4 >_{j,w} w_2$

(45) reduces to:

(46) $[w_1 >_{j,w} w_3] \wedge [w_4 >_{j,w} w_2]$

By parity of reasoning, (47) and (48) are equivalent.

(47) $[\![\text{John wants to visit Rome}]\!]^w$

(48) $[w_2 >_{j,w} w_3] \wedge [w_4 >_{j,w} w_1]$

Combining the semantics for the two desire ascriptions yields:

(49) $w_4 >_{j,w} \{w_1, w_2\} >_{j,w} w_3$

In prose, visiting Paris and Rome is more desirable to John than visiting Paris only or visiting Rome only, which is in turn more desirable to John than not visiting either city.

Meanwhile, (50) is equivalent to (51).

(50) $[\![$John wants to visit Paris and Rome$]\!]^w$ =
$\forall w' \in \text{DOX}_{j,w}$: $\text{Sim}_{w'}(\text{DOX}_{j,w} \cap \{w_4\}) >_{j,w} \text{Sim}_{w'}$ $(\text{DOX}_{j,w} \cap \{w_1, w_2, w_3\})$

(51) a. $\text{Sim}_{w_1}(\{w_4\}) >_{j,w} \text{Sim}_{w_1}(\{w_1, w_2, w_3\}) \wedge$
b. $\text{Sim}_{w_2}(\{w_4\}) >_{j,w} \text{Sim}_{w_2}(\{w_1, w_2, w_3\}) \wedge$
c. $\text{Sim}_{w_3}(\{w_4\}) >_{j,w} \text{Sim}_{w_3}(\{w_1, w_2, w_3\}) \wedge$
d. $\text{Sim}_{w_4}(\{w_4\}) >_{j,w} \text{Sim}_{w_4}(\{w_1, w_2, w_3\})$

This reduces to:

(52) a. $w_4 >_{j,w} w_1 \wedge$
b. $w_4 >_{j,w} w_2 \wedge$
c. $w_4 >_{j,w} w_3 \wedge$
d. $w_4 >_{j,w} \{w_1, w_2\}$

And (52) further reduces to:

(53) $w_4 >_{j,w} \{w_1, w_2, w_3\}$

Crucially, (49) entails (53). Thus it would appear at least for this particular example that the better-worlds approach erroneously predicts conjunction introduction.

How does the better-worlds approach fare with respect to conflicting desires? The reasoning will be similar except that now we eliminate the doxastic alternative where John visits both Paris and Rome so that John's doxastic alternatives are exhausted by:

(54) a. w_1: John visits Paris but does not visit Rome.
b. w_2: John visits Rome but does not visit Paris.
c. w_3: John visits neither Paris nor Rome.

So we have:

(55) $[\![$John wants to visit Paris$]\!]^w$ =
$\forall w' \in \text{DOX}_{j,w}$: $\text{Sim}_{w'}(\text{DOX}_{j,w} \cap \{w_1\}) >_{j,w} \text{Sim}_{w'}$ $(\text{DOX}_{j,w} \cap \{w_2, w_3\})$

This is in turn equivalent to:

(56) a. $\text{Sim}_{w_1}(\{w_1\}) >_{j,w} \text{Sim}_{w_1}(\{w_2, w_3\}) \wedge$
b. $\text{Sim}_{w_2}(\{w_1\}) >_{j,w} \text{Sim}_{w_2}(\{w_2, w_3\}) \wedge$
c. $\text{Sim}_{w_3}(\{w_1\}) >_{j,w} \text{Sim}_{w_3}(\{w_2, w_3\})$

(56) reduces to:

(57) a. $w_1 >_{j,w} \{w_2, w_3\} \wedge$
b. $w_1 >_{j,w} w_2 \wedge$
c. $w_1 >_{j,w} w_3$

(57) in turn reduces to:

(58) $w_1 >_{j,w} \{w_2, w_3\}$

By parity of reasoning (59) and (60) are equivalent.

(59) $[\![\text{John wants to visit Rome}]\!]^w$

(60) $w_2 >_{j,w} \{w_1, w_3\}$

And now note that that (58) and (60) contradict each other: w_1 is asserted to be both more and less desirable to John than w_2. Thus it would appear that, at least for this particular example, the better-worlds approach erroneously makes contradictions out of conflicting desire reports.[4]

A number of solutions have been proposed for solving the problem of conjunction introduction and conflicting desires. As recently reviewed by Phillips-Brown (to appear), most of these solutions fall into one of two major categories. One category entertains the idea that *want*-sentences do not have a fixed ordering source: preferences can be ranked in different ways. Levinson (2003); Crnič (2011) both propose versions of this idea.[5] Suppose, for example, that John's de-

[4]Crnič (2011:171) claims that the better-worlds approach allows for conflicting desires in the Paris/Rome scenario if we add in doxastic alternatives where John visits Geneva: if the most similar visiting-Paris worlds in which John does not visit Paris are visiting-Geneva worlds, and the most similar visiting-Rome worlds in which John does not visit Rome are visiting-Geneva worlds, then visiting-Paris worlds and visiting-Rome worlds are better than visiting-Geneva worlds. But this does not actually solve the problem of conflicting desires, for two reasons pointed out to me by Milo Phillips-Brown (p.c.). First, on the better-worlds approach, if *John wants to visit Paris* is true, the relevant desirability comparison has to hold at each of John's doxastic alternatives. In Crnič's scenario, among these alternatives will be worlds where he visits Rome. So the truth of the sentence will depend on Paris being preferable to Rome, and vice versa for *John wants to visit Rome*, thereby reintroducing conflicting desires. The presence of a third alternative does not help. Second, even if third alternatives did help, it is easy to imagine conflicting desires in the absence of a third alternative; that is, even if John really believes that he only has two options for his destination city, the semantics needs to allow for conflicting desires.

[5]In addition to multiple ordering sources, Levinson's (2003) account of *want*-sentences also has a probabilistic component that borrows concepts from expected utility theory. See also Lassiter 2011 for an approach similar to Levinson's but without multiple ordering sources. For Lassiter, *want*-sentences are true iff the expected value of the desired outcome is significantly greater than the average expected value of contextually supplied alternatives. According to Lassiter, this approach enables a scenario where visting Paris and visiting Rome can both be wanted even if they conflict with each other, provided there are enough alternatives entering the calculation to bring down the

sire to visit Paris is validated relative to an ordering source that places high value on studying French, whereas John's desire to visit Rome is validated relative to an ordering source that places high value on studying Italian. Then, there is no contradiction. (And although the terminology 'ordering source' suggests the best-worlds approach, this solution can also be implemented in the better-worlds approach, by countenancing multiple flavors of comparative desirability.) As Phillips-Brown (to appear) discusses, the major challenge facing this solution is that it needs to be reined in with principles telling us what ordering sources are available in any given context. Consider again the example in (61). There are scenarios compatible with (61) in which there is no way of reading (62) as being true. But if we are allowed to use an ordering source that places high value on seeing the concert, then we erroneously predict that (62) has a true reading relative to that ordering source. Of course this is not fatal for the multiple ordering sources approach, but it remains to be seen whether it can be reined in in a principled way.

(61) [Al knows that he'll see the concert only if he takes the long drive:]
Al wants to see the concert, but Al doesn't want to take the long drive. (adapted from Phillips-Brown to appear)

(62) Al wants to take the long drive.

The other category of solutions entertains the idea that *want*-sentences do not have a fixed modal base. In particular, Villalta (2008) proposes that the modal base can include worlds outside the attitude holder's doxastic alternatives (cf. also Rubinstein 2012 and section 6.8 below). The primary challenge facing this approach is similar to the one facing the multiple ordering sources approach: we are owed a principled account of which worlds outside the attitude holder's doxastic alternatives are allowed into the modal base. Phillips-Brown (2016) entertains a couple of possibilities but points out problems for both and ends up proposing an alternative account of conflicting desires whereby *want*-sentences have just one fixed ordering source and one fixed modal base, but the modal base employs what Phillips-Brown calls COARSE-WORLDS: worlds that do not decide the truth of every proposition. Phillips-Brown argues that this gives us the right kind of flexibility we need to handle conflicting desires in an appropriately constrained way.

One other recent approach to conflicting desires that deserves mention is that proposed by Condoravdi and Lauer (2016). These authors propose that an agent at a world has a set of preference structures, each structure encoding some preference source such as desire, inclination, personal moral code, and obligation. Each one of these preference structures is modeled as a set of propositions plus an importance

average expected value of the standard of comparison. This approach bears some interesting connections to the fact that *want* is gradable and focus-sensitive; see sections 6.7 and 6.8 respectively.

ranking. On their view, *want*-sentences are underspecified for preference structure and in that sense this view is similar to the one mentioned above wherein we allow multiple ordering sources. But these authors also allow for conflicting desires even with a fixed preference structure by proposing that the truth of a sentence a *wants* p simply depends on p being a member of the set of highest-importance-ranked propositions in the relevant preference structure associated with a. There are no inherent constraints on the logical relationship among these propositions, so the set can freely contain contradictory propositions. That being said, the lack of logical relationships also means that *want*-sentences end up not being monotonic. Condoravdi and Lauer suggest that apparently monotonic behavior can be analyzed by scoping the offending material outside the *want*-predicate. But it is unlikely that all apparent cases of monotonicity can be handled in this fashion: see especially Fara 2013 and section 6.5 above.

Ultimately the verdict is still out on the proper way of handling conflicting desires. Whatever the correct solution ends up being, it seems likely that it will also make sense of two other intuitively related and noteworthy properties of *want*-sentences, namely gradability and focus sensitivity, to which we now turn.

6.7 Wanting it very much: Gradability

It has occasionally been noted in the literature that unlike *believe*, *want* behaves like a gradable predicate (Levinson 2003; Villalta 2008; Lassiter 2011; Anand and Hacquard 2013; Grano 2017b). For example, the following data taken from Grano 2017b:7 show that unlike *believe*, *want* participates in comparative constructions (63-a)/(64-a) and superlatives (63-b)/(64-b), and admits degree modifiers like *very much* (63-c)/(64-c).

(63) a. John wants to go to Paris more than he wants to go to London.
b. What John wants the most is to be happy.
c. John wants very much to leave.

(64) a. ?John believes he'll go to Paris more than he believes he'll go to London.
b. ?What John believes the most is that he'll be happy.
c. ?John believes very much that he'll leave.

Neither the best-worlds approach nor the better-worlds approach to *want*-sentences are equipped to handle such data without adjustments to the proposals. In the best-worlds approach, the ordering source built into the meaning of *want* singles out the most desirable worlds, sealing off any access to less desirable worlds that would be needed to handle a comparison of relative desirability like (63-a).

The better-worlds approach, while based on a notion of comparative desirability of worlds, also falls short, since the standard of comparison is hardwired in and therefore cannot be manipulated in the way that something like (63-a) would require.

In order to handle cases like (63-a), then, what seems to be needed is to strip out from the meaning of *want* the superlativity (on the best-worlds approach) or the fixed standard of comparison (on the better-worlds approach). Villalta (2008) makes a proposal exactly along these lines. Just as gradable predicates like *tall* are standardly treated as relations between degrees and individuals, as in (65) (see e.g. Kennedy 1999 among many others), Villalta proposes to treat *want* as a relation between degrees, propositions and individuals, as in (66). This degree argument can then be manipulated by a comparative morpheme, as in (63-a), a superlative morpheme, as in (63-b), or by a degree modifier like *very much* as in (63-c).

(65) $[\![\text{tall}]\!] = \lambda d \lambda x.$x is tall to degree d

(66) $[\![\text{want}]\!] = \lambda d \lambda p \lambda x.$x wants p to degree d

What about *want*-sentences that do not involve an overt degree morpheme? Again taking guidance from the degree semantics literature, the typical move here is to propose that unmodified uses of gradable adjectives like *tall* as in (67) undergo a type-shift or combine with a silent morpheme POS(ITIVE) that interacts with the rest of the material in the sentence to yield a meaning like 'John is tall to some degree d that exceeds a contextually determined threshold.' (see e.g. Kennedy 2007).

(67) John is tall.

Villalta (2008) suggests that unmodified *want*-sentences similarly involve a silent degree morpheme, but rather than assigning it a positive semantics as is needed for gradable adjectives, she assigns it a superlative semantics that combines with *want* so as to result in (her version of) Heim's better-worlds semantics for *want*-sentences: the desired proposition is asserted to be more desirable than *any* relevant alternative (where for Heim the relevant alternatives are calculated by taking the negation of the desired proposition, though Villalta has a different take on this — see section 6.8 below).

Is it right to assign a superlative semantics to unmodified *want*-sentences? One potential reason for thinking otherwise has to do with some of the cases of conflicting desires considered above. Recall from section 6.6 above the scenario where John only has the resources to visit one city this summer but nonetheless (68-a) and (68-b) can both be true. It is difficult to see how (68-a–b) can be reconciled with a superlative semantics (unless one allows variable ordering sources) and in fact an explicitly superlative variant sounds somewhat odd, as in (69). But if *want*-sentences assert something a bit weaker — say, that going to Paris is "good

enough", then it is much easier to see how (68-a) and (68-b) needn't contradict each other. (See also Lassiter 2011 for a similar idea.) But even if this idea is right, it remains to be embedded into a more general theory of what counts as "good enough" and whether the logic of positive semantics works just the same for *want*-sentences as it does for better studied exemplars of gradability like *tall*.

(68) a. John wants to go to Paris.
b. John wants to go to Rome.

(69) ??John wants to go to Paris more than anywhere else and John wants to go to Rome more than anywhere else.

6.8 Wanting JOHN to teach syntax on Tuesdays and Thursdays: Focus sensitivity

Building on earlier observations by Dretske (1972, 1975) (cf. also von Fintel 1999), Villalta (2008) argues that yet another difference between *believe* and *want* is that *want* is focus-sensitive; i.e., the truth conditions of *want*-sentences are affected depending on what in the complement of *want* is focused.

Villalta has us consider a scenario wherein teaching schedules and teaching assignments are being discussed at a department faculty meeting. Lisa would prefer to have Lara teach syntax rather than to have John teach syntax, but given that John is the only option for teaching syntax, she would prefer that he teach syntax on a Tuesday-Thursday schedule rather on a Monday-Wednesday-Friday schedule. Then the crucial observation is that in this scenario, (70-a) is truthful because it has a meaning along the lines of (70-b), whereas (71-a) is false because it has a meaning along the lines of (71-b). (I follow Villalta here in using all caps to signal the focused constituent.)

(70) a. Lisa wants John to teach syntax ON TUESDAYS AND THURSDAYS.
b. ≈ John teaching syntax on Tuesdays and Thursdays is more desirable to Lisa than John teaching syntax on other kinds of weekly schedules.

(71) a. Lisa wants JOHN to teach syntax on Tuesdays and Thursdays.
b. ≈ John teaching syntax on Tuesdays and Thursdays is more desirable to Lisa than someone else teaching syntax on Tuesdays and Thursdays.

Villalta uses these and other facts to argue for a modified version of Heim's better-worlds semantics for *want*-sentences, given in (72).[6] Whereas Heim's semantics involves universal quantification over the attitude holder's doxastic alternatives,

[6]Aside from focus sensitivity, Villalta argues that her revision to Heim's semantics is also sup-

Villalta's revised approach involves universal quantification over a set of contextually supplied propositional alternatives (the connection to belief still achieved in that the propositional alternatives are presupposed to be consistent with the attitude holder's doxastic alternatives). And whereas Heim's semantics involves the relative desirability of p-worlds over maximally similar $\neg p$-worlds, Villalta's semantics involves the relative desirability of p-worlds over the contextually supplied alternatives. Building on Rooth's (1982; 1992) theory of focus interpretation, the computation of the alternatives is affected by focus, thereby leading to an account of the data in (70)–(71) above.

(72) $[\![\text{want}_C]\!]^g$(p)(a)(w) is defined iff $\forall$ q $\in$ g(C): $\text{Dox}_{a,w} \cap$ q $\neq \emptyset$
if defined $[\![\text{want}_C]\!]^g$(p)(a)(w) = 1 iff
$\forall$ q: q $\neq$ p $\wedge$ q $\in$ g(C): p $>_{a,w}$ q (Villalta 2008:480)

It seems highly plausible that there is a deep analytical connection between the facts from the last three subsections concerning conflicting desires, gradability, and focus sensitivity: all point toward the conclusion that *want*-sentences are more relative and more context-sensitive than *believe*-sentences. While the full range of consequences and predictions of Villalta's approach is yet to be explored, it is no doubt a promising development.

6.9 Selling your cello: Presupposition projection

One might be struck by the fact that so much of the discussion in this chapter has been framed as a reaction to the initial hypothesis that *want*-sentences express universal quantification over bouletic alternatives. Suppose we had instead started with the hypothesis that *want*-sentences express some relation between individuals and propositions while remaining neutral about the nature of that relation. Then, we would have not made any faulty predictions about entailment relations between *believe*-sentences and *want*-sentences (section 6.3), the felicity of desire ascriptions with contents believed to be inevitable or impossible (section 6.4), monotonicity (section 6.5), or conjunction introduction and conflicting desires (section 6.6). Meanwhile, gradability (section 6.7) and focus sensitivity (section 6.8) would have both still required some treatment, but not necessarily one that necessitates reference to the attitude holder's doxastic alternatives.

So is there any good reason to think that *want*-sentences involve the attitude holder's doxastic alternatives if one is not pre-committed to an underlyingly

ported by facts about *want*-sentences in scenarios where more than two alternatives are present. But see Rubinstein 2012:111–117 for a reply to this argument, showing that the data are in fact consistent with Heim's (1992) negation-based version of the better-worlds analysis.

Hintikka-like architecture for attitude reports? No doubt there are important connections between desire and belief, but why think that this is the business of grammar? To address this question we will come back full circle to Heim (1992) and come clean about her paper's framing motivation, which was not after all to make sense of teaching-Tuesdays-and-Thursdays cases but rather to make sense of some puzzling facts about presupposition projection.

In particular, Heim begins with an observation due to Karttunen (1974) that, "if the complement of an attitude sentences presupposes p, then that sentence as a whole presupposes that the attitude-holder believes p" (Heim 1992:183). For example, the sentence *Patricks sells his cello* presupposes that Patricks owns a cello. Heim observes that whereas at first glance (73) seems to carry the same presupposition, (74) reveals that the presupposition can actually merely be that Patrick *believes* that he owns a cello.

(73) Patrick wants to sell his cello.

(74) Patrick is under the misconception that he owns a cello, and he wants to sell his cello.

Why should desire reports cause ordinary presuppositions to turn into presuppositions about the attitude holder's beliefs? Heim's proposal was that this is because at some level of representation, computing the meaning of (73) involves adding the proposition named by *Patrick sells his cello* to the set of worlds constituting Patrick's beliefs, thereby inducing a presupposition about belief, exactly parallel to the way an unembedded use of *Patrick sells his cello* would be added to the set of worlds constituting the Common Ground, thereby inducing a presupposition about the Common Ground.

Much subsequent work on *want*-sentences, although largely building on Heim's approach, has been motivated by concerns other than presupposition projection; von Fintel (1999), for example, was interested in the distribution of negative polarity items, and Villalta (2008) was interested in the relative distribution of indicative and subjunctive mood in Spanish. Geurts (1998) and Maier (2015) both deal with presupposition projection in *want*-sentences and both follow Heim in making important use of the idea that *want*-sentences have a doxastic component. Maier in particular, in his appropriately titled paper "Parasitic attitudes", proposes that "desires and other non-doxastic attitudes are asymmetrically dependent on beliefs" (p. 205).

To be sure, the behavior of *want*-sentences with respect to presupposition projection does not *prove* that *want*-sentences have a doxastic component. But it does constitute grammatical evidence in favor of such a connection, and one that does not depend on any special starting assumptions about how attitude reports are

supposed to work. Anyone who proposes that *want*-sentences simply name some unspecified relation between individuals and propositions owes us an account of the observed presupposition facts.

6.10 Scaling up: The typology of attitude predicates

In (75) we take stock of the empirical differences between *want* and *believe* that have been the focus of this chapter.

(75)
a. *want* is parasitic on belief semantics; *believe* is not parasitic on desire semantics.
b. *want* gives rise to (apparently) non-monotonic behavior; *believe* does not.
c. *want* (apparently) does not obey conjunction introduction; *believe* does.
d. Paired *want*-sentences can have conflicting contents without portraying the attitude holder as irrational; paired *believe*-sentences cannot.
e. *want* is gradable; *believe* is not.
f. *want* is focus-sensitive; *believe* is not.

Should we pursue a theory in which all of these differences fall out from minimal starting assumptions? Or should we pursue a more modular approach whereby some or all of these differences are independent from one another? The answer to that question will depend in no small part on the extent to which the observed clustering of properties is accidental or principled. And to investigate this question, we need to scale our investigation up to other attitude predicates both in English and in other languages. Indeed, one of the major interests in probing the fine-grained semantic differences between *believe* and *want* lies not just in better understanding the semantics of these two verbs *per se* but also in using them as a starting point for investigating semantic uniformity and variation across the full range of attitude predicates that are attested in natural language.

In this connection, Stalnaker (1984) adopts the term *acceptance* as a "broader concept than belief; it is a generic propositional attitude concept with such notions as presupposing, presuming, postulating, positing, assuming and supposing as well as believing falling under it" (p. 79). Stalnaker goes on to say that, "To accept a proposition is to treat it as a true proposition in one way or another — to ignore, for the moment at least, the possibility that it is false" (p. 79). As a diagnostic, he suggests that "one may say that a propositional attitude concept is an acceptance concept if the attitude is said to be *correct* whenever the proposition is true" (p. 79–80). On the linguistic end, we might correspondingly follow Stalnaker's lead in

hypothesizing that the verbs *believe*, *presuppose*, *presume*, *postulate*, *posit*, *assume*, and *suppose* all name attitudes of acceptance. And if they all pattern as *believe* does with respect to the properties enumerated in (75), that would constitute support for the view that this particular clustering of properties is not accidental.

Two natural follow-up questions are: is *want* also a member of a broader family of attitude predicates that share the same basic properties? And are there yet other attitude predicates that cannot be neatly categorized as *believe*-like or *want*-like? In the remainder of this section we'll take up these two questions in turn.

Turning first to the question of *want*'s kinship, I first want to lay bare an obvious syntactic fact that we have been ignoring up to this point: *believe* can combine with finite clauses whereas *want* cannot:

(76) a. Beatrix believes that it will rain tomorrow.
b. *Beatrix wants that it will rain tomorrow.

Could this syntactic difference be a correlate of one or more of the semantic differences we have uncovered? Consider also Stalnaker's proposal that belief is an attitude of acceptance — and what could be more acceptance-like than an unembedded finite clause, whose canonical function is to present a proposition as being true? This parallelism between the form of unembedded assertions and the form of complements to *believe* was not lost on Stalnaker, who quoted Fodor's position that "it could hardly be an accident that declarative sentences of English constitute the (syntactic) objects of verbs like 'believe'" (Fodor 1978:503, quoted by Stalnaker 1984:59).

What this suggests is that we might sort attitude predicates into two classes depending on whether they can combine with a finite clause, and then investigate the semantic concomitants of the two classes. As it turns out, the relative distribution of finite and nonfinite clauses in English ends up being somewhat messy. Much of the mess is cleaned up by recognizing that there are several kinds of nonfinite clauses: one must distinguish (at least) raising infinitives, control infinitives, *for-to* infinitives, and ECM or raising-to-object infinitives. With these finer-grained categories, interesting semantic correlates are discoverable; see e.g. Moulton 2009 who argues that ECM infinitives occur only in doxastic contexts and Grano 2016 (cf. also Portner 1997) who argues that *for-to* infinitives are essentially the opposite, occurring only in non-doxastic contexts.

But if we turn to the Romance languages, we find a similar phenomenon with a somewhat cleaner distribution and whose semantic correlates have been investigated much more extensively: mood choice. In Romance, finite complement clauses occur with either indicative or subjunctive mood morphology, and the choice is determined in part by the embedding verb. With some important excep-

tions, translation equivalents of 'believe' tend to combine with indicative complements whereas translation equivalents of 'want' combine with subjunctive complements. And the facts that emerge when we scale up from 'believe' and 'want' render plausible the hypothesis that mood choice is not idiosyncratic but rather reflects some semantic properties of the embedding verb. Based in part on mood choice facts, Bolinger (1968) proposes a distinction between embedding predicates that are *representational* (strikingly similar to Stalnaker's concept *acceptance*), which take the indicative, and embedding predicates that are *volitional*, which take the subjunctive. Much subsequent work has been concerned with refining these two categories in light of more recent developments in attitude semantics. Giorgi and Pianesi (1997), for example, working with a best-worlds semantics for *want*-sentences, propose that the predicates that select for the subjunctive are those that have a non-null ordering source. In a different vein, Villalta (2008), working with her own variant of the better-worlds semantics for *want*-sentences, proposes that in Spanish, the predicates that select for the subjunctive are those that are gradable. These include translation equivalents of *want*, *prefer*, *fear*, *regret*, *be glad*, *be surprised*, *doubt*, *order*, *advise*, *suggest*, *make*, and *achieve*. Another influential approach, exemplified by Giannakidou (1998, 1999), proposes that the indicative mood is selected by veridical attitude predicates; veridicality is a concept bearing close similarity to acceptance in Stalnaker's sense and representationality in Bolinger's sense.

Aside from mood choice, another phenomenon that has been found to be relevant in identifying categories of attitude predicates has to do with the availability of epistemic interpretations for embedded modals. Anand and Hacquard (2013), in particular, show that the modal verb *have* (which is in principle compatible with both deontic and epistemic interpretations) can have an epistemic interpretation when embedded under *believe*, as in (77), but not when it is embedded under *want*, as in (78). Generalizing out from here, the authors propose that epistemic modals are available only under attitudes of acceptance in Stalnaker's sense, and they propose an account whereby attitudes of acceptance are those attitudes that quantify over an information state, to which epistemic modals bear an anaphoric dependency.

(77) John believes that Paul has to be innocent.

(78) John wants Paul to have to be innocent.

We turn now to the second question raised above: are there attitude predicates that cannot be neatly categorized as *believe*-like or *want*-like? In a word, yes. One salient example is *hope* (and its negative counterpart *fear*): as discussed by Anand and Hacquard (2013), *hope* is *believe*-like in that it asserts something about the attitude holder's doxastic state: a *hopes* p entails that a believes p to be

possible. Compare (79), repeated from above, which illustrates that *want* tolerates attitude holder-relative counterfactuality, with (80).

(79) I want this weekend to last forever. (But I know, of course, that it will be over in a few hours.) (Heim 1992:199)

(80) #I hope this weekend will last forever. (But I know, of course, that it will be over in a few hours.)

At the same time, *hope* is *want*-like in that a *hopes* p also asserts that a has a preference for p. Given these mixed semantic properties it is surely no accident that it also displays mixed mood selection behavior: in Spanish, for example, *esperar* 'hope' selects the subjunctive whereas in French, *espérer* 'hope' selects the indicative (and see Portner and Rubinstein 2012 for a relevant account of mood choice in French). In fact, Anand and Hacquard (2013) show that *hope* is even mixed with respect to the embedded epistemic modal diagnostic: it allows embedded epistemic possibility modals but not embedded epistemic necessity modals.

Another class of attitude predicates that warrant discussion are those that involve commitment to action. Grano (2017b), for example, drawing in part on Sag and Pollard (1991), identifies predicates that involve an internal mental commitment to take action such as *decide*, *intend*, *plan*, and *try*; predicates that involve a public commitment to take action such as *agree*, *offer*, *promise* and *swear*; and predicates that involve a cause or attempted cause to create a commitment to take action such as *beg*, *order*, *persuade*, and *urge*. (See also White and Rawlins to appear on *decide* and Condoravdi and Lauer 2011 on *promise*.)

Focusing on *intend*, Grano (2017b) shows that *intend* patterns in some ways like *believe*, in other ways like *want*, and in yet other ways unlike both. Grano argues that the *believe*-like properties of *intend* follow from the fact that *intend* involves commitment to action, which endows it with Hintikka-esque rational properties not found with *want*. The *want*-like properties, Grano argues, follow from the fact that *want* and *intend* are both based on a preference semantics (this builds on Condoravdi and Lauer 2016). And the properties unique to *intend* stem from the fact that *intend*, unlike both *believe* and *want*, establishes a relation of responsibility between the attitude holder and the outcome named by its complement (this builds on Farkas 1988).

The predicate *try* also deserves special mention: unlike *believe*, *want* and *intend*, *try* is used not only to name an attitude but also to assert that action has been undertaken. Sharvit (2003) argues that this action component of *try* renders it highly similar to modal analyses of progressive aspect; see also Grano (2011, 2017a).

It should be apparent from this brief survey that there is a wide range of semantic variation in attitude predicates, and still a lot of work to be done on the less-studied ones. The fine-grained similarities and differences that have been uncovered so far suggest that ultimately, the question is not "how many categories of attitude predicates are there?" but rather "what are the semantic parameters along which attitude predicates divide and what constraints are there on combinations of parameter settings?"

6.11 Further reading

Three essential starting points for *want*-sentences are: Stalnaker 1984: Chapter 5; Heim 1992; von Fintel 1999. Subsequent work that deserve close study include: Levinson 2003; Villalta 2008; Crnič 2011: Chapter 3 and Appendix; Lassiter 2011: Chapters 5–6; Rubinstein 2012: Chapter 3; Condoravdi and Lauer 2016; Phillips-Brown to appear. For a more philosophically oriented perspective, see Fara 2013 and Braun's (2015) reply. As for the broader typology of attitude predicates, Moltmann 1994; Asher 1987 are useful resources. The vast literature on mood choice is also an excellent point of entry; see Portner 2018 for a recent survey.

6.12 Discussion questions

1. Consider the following hypothetical exchange:

   (81) a. The sun is going to rise at 6:46 a.m. tomorrow.
   b. Good! I want the sun to rise at 6:46 a.m. tomorrow!

   Assuming that this is a felicitous exchange, how does it bear on Heim's suggestion that *want*-sentences have an undefined semantics when the prejacent is entailed by the attitude holder's doxastic alternatives (or a superset thereof that ignores beliefs about intended actions)?

2. As discussed above, Heim's better-worlds semantics for *want*-sentences erroneously predicts (before adding the presuppositional component) that if a *believes* p or a *believes* $\neg p$ is true, then a *wants* p is also true. von Fintel's best-worlds approach also erroneously predicts an entailment from a *believes* p to a *wants* p. But does it also predict an entailment from a *believes* $\neg p$ to a *wants* p? Why or why not?

3. Whereas *want* cannot take an explicit standard of comparison unless comparative morphology is added, neighboring expressions *would rather* and *prefer*

both do so, in the former case introduced by *than* (82-a) and in the latter case introduced by *to* (82-b). Can the technology discussed in this chapter be recruited to assign an appropriate semantics to sentences like these? What would an appropriate semantics look like? And do expressions like these lend credence to the idea that the grammar treats desirability as an inherently gradable notion (unlike, say, belief)?

(82) a. John would rather visit Paris than Rome.
b. John prefers visiting Paris to visiting Rome.

4. Farkas (1992) coins the term FICTION verbs for a class of predicates including *dream*, *imagine*, and *lie* which are interesting in that they can combine with finite clauses (and Romance equivalents combine with indicative complements) despite the fact that they are ordinarily used in contexts where neither the speaker nor the attitude holder seems to consider the content of the attitude true. Should these verbs be considered attitudes of acceptance in Stalnaker's (1984) sense? Why or why not? And if so, does Stalnaker's diagnostic for what counts as an attitude of acceptance need to be revised at all?

References

Anand, Pranav, and Valentine Hacquard. 2013. Epistemics and attitudes. *Semantics & Pragmatics* 6:1–59.

Asher, Nicholas. 1987. A typology for attitude verbs and their anaphoric properties. *Linguistics and Philosophy* 10:125–197.

Bolinger, Dwight. 1968. Post-posed main phrases: An English rule for the Romance subjunctive. *Canadian Journal of Linguistics* 14:3–30.

Bratman, Michael E. 1987. *Intentions, plans, and practical reason*. Cambridge, MA: Harvard University Press.

Braun, Davis. 2015. Desiring, desires, and desire ascriptions. *Philosophical studies* 172:141–162.

Condoravdi, Cleo. 2002. Temporal interpretation of modals: Modals for the present and for the past. In *The construction of meaning*, ed. David Beaver, Stefan Kaufmann, Brady Clark, and Luis Casillas, 59–88. Stanford: CSLI Publications.

Condoravdi, Cleo, and Sven Lauer. 2011. Performative verbs and performative acts. In *Sinn und Bedeutung 15*, ed. Ingo Reich, Eva Horch, and Dennis Pauly, 149–164. Saarbrüken: Universaar.

Condoravdi, Cleo, and Sven Lauer. 2016. Anankastic conditionals are just conditionals. *Semantics & Pragmatics* 9:1–61.

Crnič, Luka. 2011. Getting *even*. Ph.D. Dissertation, Massachusetts Institute of Technology.

Dretske, Fred. 1972. Contrastive statements. *Philosophical Review* 81:411–437.

Dretske, Fred. 1975. The content of knowledge. In *Forms of representation*, ed. Bruce Freed, A. Marras, and P. Maynard, 77–93. Amsterdam: North Holland Publishing Company.

Fara, Delia Graff. 2013. Specifying desires. *Noûs* 47:250–272.

Farkas, Donka. 1992. On the semantics of subjunctive complements. In *Romance languages and modern linguistic theory*, ed. P. Hirschbueler and K. Koerner, 69–104. Amsterdam and Philadelphia: Benjamins.

Farkas, Donka F. 1988. On obligatory control. *Linguistics and Philosophy* 11:27–58.

von Fintel, Kai. 1999. NPI-licensing, Strawson-entailment, and context-dependency. *Journal of Semantics* 16:97–148.

Fodor, Jerry A. 1978. Propositional attitudes. *The Monist* 61:501–523.

Geurts, Bart. 1998. Presuppositions and anaphors in attitude contexts. *Linguistics and Philosophy* 21:545–601.

Giannakidou, Anastasia. 1998. *Polarity sensitivity as (non)veridical dependency*. Amsterdam: John Benjamins Publishing Company.

Giannakidou, Anastasia. 1999. Affective dependencies. *Linguistics and Philosophy* 22:367–421.

Giorgi, Alessandra, and Fabio Pianesi. 1997. *Tense and aspect*. New York: Oxford University Press.

Grano, Thomas. 2011. Mental action and event structure in the semantics of *try*. In *Proceedings of SALT 21*, ed. Neil Ashton, Anca Chereches, and David Lutz, 426–443. eLanguage.

Grano, Thomas. 2016. A coercion-free semantics for intention reports. In *Proceedings of the fifty-first annual meeting of the Chicago Linguistic Society*, ed. Ksenia Ershova, Joshua Falk, Jeffrey Geiger, Zachary Hebert, Robert E. Lewis Jr., Patrick Munoz, Jacob B. Phillips, and Betsy Pillion, 213–223. Chicago, IL: Chicago Linguistic Society.

Grano, Thomas. 2017a. Control, temporal orientation, and the cross-linguistic grammar of *trying*. *Glossa* ?:pp–pp.

Grano, Thomas. 2017b. The logic of intention reports. *Journal of Semantics.* https://doi.org/10.1093/jos/ffx010.

Heim, Irene. 1992. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9:183–221.

Karttunen, Lauri. 1974. Presupposition and linguistic context. *Theoretical Linguistics* 1:181–194.

Kennedy, Christopher. 1999. *Projecting the adjective: The syntax and semantics of gradability and comparison*. New York: Garland.

Kennedy, Christopher. 2007. Vagueness and grammar: the semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy* 30:1–45.

Kratzer, Angelika. 1981. The notional category of modality. In *Words, Worlds, and Contexts*, ed. J. Eikmeyer and H. Riesner, 38–74. Berlin: Walter de Gruyter.

Lassiter, Daniel. 2011. Measurement and modality. PhD dissertation, New York University.

Levinson, Dmitry. 2003. Probabilistic model-theoretic semantics for *want*. In *Proceedings of SALT 13*, ed. R. Young and Y. Zhou, 222–239. Ithaca, NY: Cornell University.

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

Maier, Ernst. 2015. Parasitic attitudes. *Linguistics and Philosophy* 38:205–236.

Moltmann, Friederike. 1994. Attitude reports, events, and partial models. Ms., CUNY Graduate Center.

Moulton, Keir. 2009. Natural selection and the syntax of clausal complementation. PhD dissertation, University of Massachusetts.

Phillips-Brown, Milo. 2016. I want to, but.... Handout from talk given at Sinn und Bedeutung 21 in Edinburgh, September 2016.

Phillips-Brown, Milo. to appear. I want to, but.... To appear in *Proceedings of Sinn und Bedeutung* 21.

Portner, Paul. 1997. The semantics of mood, complementation, and conversational force. *Natural Language Semantics* 5:167–212.

Portner, Paul. 2005. *What is meaning? Fundamentals of formal semantics*. Malden, MA: Blackwell Publishing.

Portner, Paul. 2018. *Mood*. New York: Oxford University Press.

Portner, Paul, and Aynat Rubinstein. 2012. Mood and contextual commitment. In *The proceedings of SALT 22*, ed. Anca Chereches, 461–487. Ithaca, NY: CLC Publications.

Rooth, Mats. 1982. Association with focus. Ph.D. dissertation, University of Massachusetts, Amherst.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1:75–116.

Ross, Alf. 1941. Imperatives and logic. *Theoria* 7:53–71.

Rubinstein, Aynat. 2012. Root modalities and attitude predicates. Ph.D. Dissertation, University of Massachusetts Amherst.

Sag, Ivan, and Carl Pollard. 1991. An integrated theory of complement control. *Language* 67:63–113.

Sharvit, Yael. 2003. Trying to be progressive: The extensionality of *try*. *Journal of Semantics* 20:403–445.

Stalnaker, Robert. 1968. A theory of conditionals. In *Studies in logical theory*, ed. N. Resher, 98–112. Oxford: Blackwell.

Stalnaker, Robert. 1984. *Inquiry*. Cambridge: MIT Press.

Villalta, Elisabeth. 2008. Mood and gradability: An investigation of the subjunctive mood in Spanish. *Linguistics and Philosophy* 31:467–522.

White, Aaron Steven, and Kyle Rawlins. to appear. Question agnosticism and change of state. To appear in *Proceedings of Sinn und Bedeutung* 21.