

Multidimensionality, subjectivity and scales: experimental evidence

Stephanie Solt

Zentrum für Allgemeine Sprachwissenschaft (ZAS), Berlin

Abstract

This paper investigates the subjective interpretation of the comparative form of certain gradable adjectives, exploring in particular the hypothesis put forward in several recent works that such ‘ordering subjectivity’ derives from the multidimensional nature of the adjectives in question. Results are presented of an experimental study showing that ordering subjectivity is more widespread than previously recognized, and that in this respect, gradable adjectives divide into not two but three groups: objective, subjective and mixed. Some issues are discussed with defining the class of multidimensional adjectives, and on the basis of these observations as well as the experimental findings, it is argued that there are two separate sources of ordering subjectivity: multidimensionality and judge dependence. A theory of gradable adjective meaning is outlined according to which the availability of subjective versus objective readings for the comparative is determined by the formal properties of adjectival measure functions.

1 Introduction

It is well known that certain adjectival predicates are subjective or ‘judge-dependent’, in that two competent speakers can disagree as to whether the predicate applies, without either appearing to have said something incorrect or false. Such ‘faultless disagreement’ is observed most classically with so-called predicates of personal taste such as *tasty* and *fun* (1a), but also with evaluative adjectives more generally (e.g. *beautiful*) in (1b)) and with the positive forms of vague gradable adjectives (e.g. *tall* in (1c)):

- (1) a. A: The chili is tasty! **faultless**
 B: No, it’s not tasty at all!
- b. A: The Picasso is beautiful! **faultless**
 B: No, it’s ugly!
- c. A: Anna is tall! **(potentially) faultless**
 B: No, she’s not!

Recently, attention has turned to a second sort of subjectivity, which is observed with the comparative form of some but not all gradable adjectives (Kennedy 2013; Bylinina 2014; Umbach to appear; McNally and Stojanovic 2015). For example, two competent speakers might faultlessly disagree as to which of two dishes is tastier (2a), or which of

two paintings is more beautiful (2b), but not about which of two individuals is taller (2c). In what follows, I will refer to the phenomenon exemplified in (2a-b) as **ordering subjectivity**.

- (2) a. A: The chili is tastier than the soup! **faultless**
 B: No, soup is tastier!
- b. A: The Picasso is more beautiful than the Miró. **faultless**
 B: No, the Miró is more beautiful.
- c. A: Anna is taller than Zoe. **factual only**
 B: No, Zoe is the taller of the two!

For the leading semantic approach to gradability, namely the degree based approach of Cresswell (1977); Kennedy (1997); Heim (2000) and others, ordering subjectivity is problematic. In such a framework, the meaning of gradable adjectives is taken to be based on measure functions that map individuals to degrees on scale: *tall* is based on a height measure function, *beautiful* on a beauty function, and so forth (3). Comparison constructions are then analyzed as expressing a relation between the degrees corresponding to the two compared individuals (4).

- (3) a. $\llbracket \text{tall} \rrbracket = \lambda d \lambda x. \mu_{\text{HEIGHT}}(x) \succeq d$
 b. $\llbracket \text{beautiful} \rrbracket = \lambda d \lambda x. \mu_{\text{BEAUTY}}(x) \succeq d$
- (4) The Picasso is more beautiful than the Miró.
 $\mu_{\text{BEAUTY}}(\text{Picasso}) \succ \mu_{\text{BEAUTY}}(\text{Miro})$

The mostly unspoken assumption underlying lexical entries of this form is that each dimension of measurement *DIM* is uniquely associated with a measure function μ_{DIM} whose output encodes the ordering of individuals relative to *DIM*. But examples such as (2a,b) suggest that this can't be right. Rather, it seems that measure functions must in some way be relativized to speakers, thereby allowing disagreement as to orderings.

The objective of this paper is to work towards an account of ordering subjectivity within a degree-based semantic framework. In particular, I will investigate the hypothesis, put forth in several recent works, that a – or the – source of ordering subjectivity is the **multidimensionality** of the predicates in question (Kennedy 2013; Bylinina 2014; Umbach to appear; McNally and Stojanovic 2015). Whereas the attribution of a predicate such as *tall* is based on a single underlying dimension, namely height, that of a predicate such as *beautiful* is based on multiple underlying component dimensions; for (1b) and (2b), for example, the component dimensions of beauty might involve line, color, draftsmanship and so forth. Subjectivity is claimed to arise due to uncertainty about or underspecification in the component dimensions and their relative weights; different choices of dimensions and weights results in different orderings, possibly reversing the relative positions of two individuals. Specifically, we will consider questions including the following: How precisely should the notion of multidimensionality be understood, and on what basis should an adjective be classified as multidimensional? Are there different sorts of multidimensionality? Are all gradable adjectives that exhibit ordering subjectivity (or subjectivity more generally) multidimensional? And what is the relationship between multidimensionality and the commonly invoked notion of judge sensitivity?

A crucial step in pursuing these questions is to clarify which gradable adjectives allow subjective interpretations of their comparative forms. For dimensional adjectives such as *tall* and evaluative adjectives such as *beautiful* and *tasty*, the picture seems clear: in the former case, statements about orderings are objective, while in the latter, they are necessarily subjective. But this is far from exhausting the broad and varied spectrum of gradable adjectives. Of particular interest are adjectives such as *clean/dirty*, *smooth/rough* and *sharp/dull*. These differ from members of the *tall* class in that they lack commonly used measurement units. But they also differ from adjectives such as *beautiful* and *tasty* in that they appear to describe physical properties of objects in the world, rather than internalized experiences. Can two speakers faultlessly disagree about which of two shirts is dirtier? which of two surfaces is rougher? which of two knives is sharper? The answers to such questions will shed light on the nature and source of ordering subjectivity, but intuitions here are shaky.

In the present paper, the results of an experimental investigation are presented which demonstrates that ordering subjectivity is far more widespread than has previously been recognized, and furthermore that in this respect gradable adjectives pattern into not two but three distinct subclasses. I will argue that these findings are best captured by positing two distinct sources of ordering subjectivity: multidimensionality and judge dependence. The structure of the paper is as follows: Section 2 presents the experiment and discusses some related phenomena. Section 3 briefly reviews past proposals for the analysis of subjectivity, with a view to assessing how well these account for the experimental findings. Section 4 delves into the issue of multidimensionality, and in particular the possibility of developing diagnostics for the class of multidimensional adjectives. In Section 5, a theory of gradable adjective meaning is outlined according to which the availability of objective versus subjective readings for the comparative is determined by formal properties of the measure functions lexicalized by the adjective. Finally, Section 6 concludes.

2 Experiment: Faultless Disagreement Paradigm

The present study aims to establish a firmer empirical basis for theoretical work by using a novel faultless disagreement paradigm to diagnose the presence of ordering subjectivity among a wide range of adjective types.

2.1 Methodology and stimuli

The experiment employed a forced choice task in which participants saw brief dialogues of the form in (5)-(7), and were asked to classify the nature of the disagreement between the two speakers. Two response options were given: “only one can be right; the other must be wrong” and “it’s a matter of opinion”. The first of these was classified as a judgment of ‘fact’, the second as a judgment of ‘opinion’.

- (5) A: John and Fred look similar but John is taller than Fred.
B: No, Fred is the taller one of the two.
- (6) A: Look – Tommy’s shirt is dirtier than the one his little brother Billy is wearing.
B: No, Billy’s shirt is dirtier than Tommy’s.

- (7) A: The necklace Susan is wearing today is uglier than the one she had on yesterday.
B: No, the one she was wearing yesterday was uglier.

A total of 35 gradable adjectives were tested; these were divided into the following categories according to their status as dimensional versus evaluative, as well as the nature of the standard for the positive form of the adjective, per Kennedy and McNally (2005)¹:

- Dimensional gradable adjectives, more specifically relative gradable adjectives with numerical measures (**RelNum**): *tall, short, old, new, expensive*
- Relative gradable adjectives without numerical measures (**RelNo**): *sharp, dull, dark, light, hard, soft*
- Absolute gradable adjectives with totally closed scales (**AbsTot**): *full, empty*
- Absolute gradable adjectives with partially closed scales (**AbsPart**): *wet, dry, straight, curved, rough, smooth, clean, dirty, salty*
- Adjectives which may be classified as evaluative in a broad sense (**Eval**): *good, bad, beautiful, pretty, ugly, easy, interesting, boring, tasty, fun, intelligent, happy, sad*

Adjectives were assigned to these categories according to judgments reported in the literature as well as linguistic tests. An adjective was classified as dimensional if it has a corresponding measurement system, which in turn was diagnosed by the possibility of modification of the comparative form by a measure phrase.² The evaluative category includes adjectives of the sort discussed in the literature under the terms ‘evaluative’ or ‘predicates of personal taste’. This is a mixed class, encompassing value, taste and aesthetic judgments, emotion words, and psychological predicates; they are united, and distinguished from the other four categories, in that they do not denote external physical properties.

The experiment was administered online via Amazon MTurk, with test items split across 4 lists. Each list contained 8-12 test items and 12 fillers. Some adjectives appeared on more than one list, in different item contexts. Fillers were split equally between two types: i) those expected to yield ‘opinion’ judgments, including vague nominal predicates (e.g. *jerk*), deontic and epistemic modals, statements of likelihood, and moral statements; ii) those expected to yield ‘fact’ judgments, based on factual statements (example: A: The judge found Frank guilty. B: No, the judge found Frank innocent.). Sample size was 20-25 per list, for a total sample size of $n = 91$. Full stimuli are available at <http://www.zas.gwz-berlin.de/fileadmin/mitarbeiter/solt/fault.pdf>.

¹The correlation between subjectivity and standard type is an intriguing one, which due to space constraints I cannot explore here. I refer the reader to Solt (2016) for discussion of the source of this link.

²The test was based on the comparative because, as is well known, many adjectives compose with measure phrases in the comparative but do not allow direct measure phrases (e.g. *ten dollars more expensive* vs. **ten dollars expensive*), and as such the comparative provides a better test for the existence of a numerical measurement system. The reasons for the restrictions on direct measure phrases are complex and seemingly idiosyncratic; see Schwarzschild (2005); Sassoon (2010) for discussion.

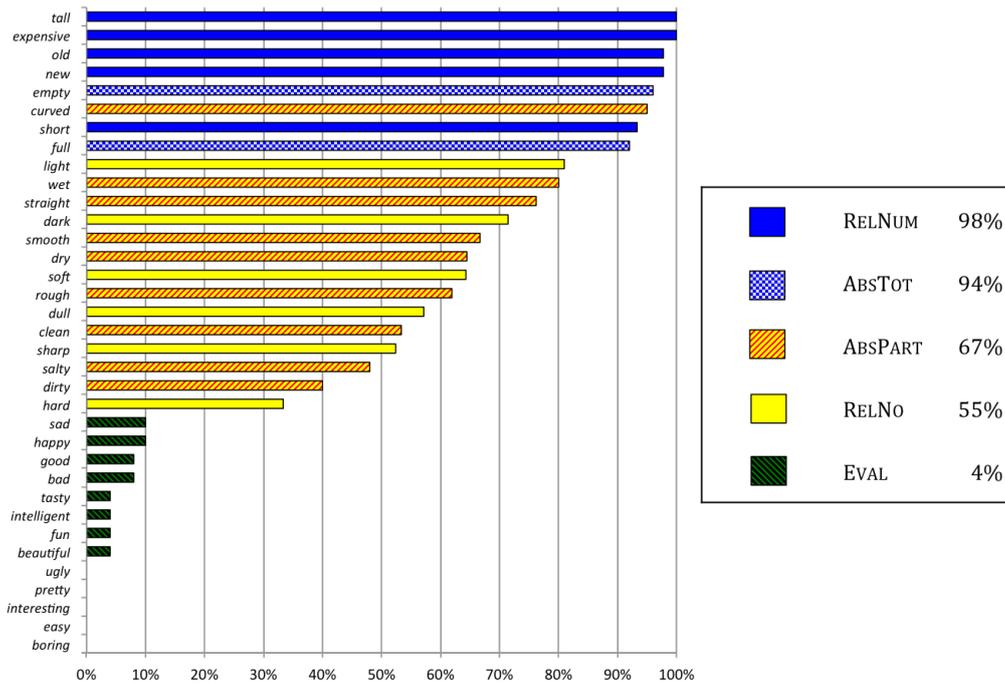


Figure 1: Results of Experiment - Percent ‘Fact’ Judgments

2.2 Results

Results by adjective class and for individual adjectives are displayed in Figure 1. A generalized linear mixed effects model was fitted to the results using the lme4 package (Bates et al. 2014) in R (R Core Team 2015), with response (‘fact’ vs. ‘opinion’) as dependent variable, adjective type as fixed effect, and random intercept for subject. The effect of adjective class was found to be significant, with post hoc testing via the multcomp package (Hothorn et al. 2008) using Tukey correction for multiple comparison showing all pairwise comparisons to be significant at $p < 0.001$ except RELNO vs. ABSPART ($p < 0.01$) and RELNUM vs. ABSTOT (nonsignificant).

2.3 Discussion and further observations

With regards to adjectives of the *tall* and *beautiful* classes, our findings are as predicted. For *tall* and the other adjectives tested that have corresponding numerical measurement systems, subjects almost universally judged disagreements about comparative statements to be factual in nature. Note that the absolute totally closed scale pair *full/empty* might be assimilated to this group, in that degrees of fullness (or emptiness) can be quantified in percentage (e.g. *90 % full, three quarters empty*). Conversely, for *beautiful, tasty* and other adjectives that were classified as evaluative, disagreements about orderings are almost universally judged to be matters of opinion.

The more interesting finding is the existence of a large group of adjectives with mixed behavior, eliciting both ‘fact’ and ‘opinion’ judgments. This group includes in particular relative gradable adjectives without corresponding measurement systems, as well as absolute gradable adjectives with partially closed scales. Among this group, we observe a range

from those adjectives that skew more towards factual readings (e.g. *straight/curved*) to those that skew towards faultless readings (e.g. *clean/dirty, salty*).

With respect to ordering subjectivity, we thus find that gradable adjectives divide into not two but rather three groups. As a caveat, it is possible that further research might determine that these groups are not as distinct as they appear to be here, or that the borders between them are not precisely where the present experiment shows them to be. That is, we cannot at this point rule out the possibility that adjectives in the ‘purely objective’ group might in certain contexts allow subjective interpretations for their comparative forms, or conversely that members of the ‘purely subjective’ class might in the right sort of contexts allow objective readings. However, one previously unrecognized finding seems relatively clear: there is a large group of adjectives whose interpretations are neither purely objective nor purely subjective.

Interestingly, the three-way division that emerges on the basis of the present faultless disagreement test is echoed in other phenomena. One of these involves measurability. Adjectives in the objective group have corresponding measurement units (in fact, the RELNUM group was defined as such). Those in the subjective group almost universally lack such units, and furthermore, for adjectives such as *fun, tasty, interesting/boring* and *beautiful/ugly*, it is hard to imagine how such units could be created (an exception in this group perhaps being *intelligent*, depending on whether one is willing to accept IQ points as a true measure of intelligence). Finally, adjectives in the mixed group fall somewhere in between. They too largely lack measurement units, but for adjectives such as *hard/soft, dark/light* and *clean/dry*, I think one has the intuition that it might be possible (say, in a laboratory setting) to establish such units. In fact, one pair in this group, namely *straight/curved*, arguably has units (*5 degrees straighter/more curved*), and perhaps not coincidentally, these pattern at the objective end of the mixed group.

A related phenomenon involves proportional comparisons. As discussed by Sassoon (2010), both dimensional and evaluative adjectives allow modification by proportional expressions such as *twice as*, and this extends also to members of the intermediate group as well (see (8)-(10)). But when we turn to precise expressions of proportion, the picture changes (see (11)-(13): these are possible for dimensional adjectives, and quite comically infelicitous for members of the evaluative class; for the intermediate group they seem marginally possible, when we imagine we are in a situation (again, say, a lab) where the dimension in question is precisely measured:

- (8) a. The Eiffel Tower is twice as tall as the Great Pyramid.
b. The laptop is five times as expensive as the tablet.
- (9) a. The Serta mattress is twice as hard as the Sealy mattress.
b. The first shirt is five times as dirty as the second one.
- (10) a. Anna is twice as beautiful as Zoe.
b. The roller coaster was ten times as fun as the ferris wheel.
- (11) a. The Eiffel Tower is 2.05 times as tall as the Great Pyramid.
b. The laptop is 4.9 times as expensive as the tablet.
- (12) a. ? The Serta mattress is 1.9 times as hard as the Sealy mattress.
b. ? The first shirt is 5.1 times as dirty as the second one.

- (13) a. # Anna is 2.3 times as beautiful as Zoe.
 b. # The roller coaster was 9.8 times as fun as the ferris wheel.

Thus the pattern observed with respect to interpretation of the comparative form appears to be part of a broader set of facts that relates to the possibility of precise, quantitative measurement.

3 Theories of subjectivity

In this section, I briefly review existing work on the semantics of subjectivity, with a view to assessing how previous proposals might account for our experimental findings.

The most extensive body of work on this topic has focused on predicates of personal taste such as *tasty* and *fun*, which provide the best-known examples of faultless disagreement: when two speakers disagree as to whether or not a dish is tasty or an experience is fun, the disagreement appears to be a real one, yet neither speaker seems to have said something incorrect or false (see Lasersohn 2005; Stephenson 2007 and other work cited below). The standard semantic approach to subjectivity of this sort is to relativize the interpretation of the adjective to a judge whose opinion or evaluation is expressed. Typically the judge is the speaker, but in certain situations this may be shifted via grammatical means or contextually to another individual. There are two classes of theories for how dependence on a judge is linguistically encoded. The relativist analysis (Lasersohn 2005) includes a judge parameter to the index of interpretation, along with the usual time and world parameters (14a). The contextualist approach (Stojanovic 2007; Sæbø 2009), by contrast, assumes that predicates of this sort feature an additional judge or experiencer argument (14b). Elaborations on and combinations of these two approaches are found in Stephenson (2007), and Bylinina (2014), among many others.

- (14) a. $\llbracket \text{tasty} \rrbracket^{w,t,j} = \lambda x.x \text{ tastes good to } j \text{ in } w \text{ at } t$
 b. $\llbracket \text{tasty} \rrbracket^{w,t} = \lambda y \lambda x.x \text{ tastes good to } y \text{ in } w \text{ at } t$

Considering these proposals from the perspective of the present topic, and in light the experimental findings, several observations can be made. First, neither of the logical forms in (14) accounts for ordering subjectivity. Although *tasty* is gradable, the above analyses localize subjectivity at the level of its positive form, thus providing no explanation for subjective judgments regarding scalar orderings. This might however be remedied fairly simply, by starting with a gradable entry of the form in (3) and relativizing the measure function – in the appropriate cases – to a judge.

A more difficult issue is that the above analyses do not provide an explanation for the finding that adjectives exhibiting ordering subjectivity divide into two groups, depending on whether or not they also allow factual readings for the comparative. If subjective adjectives are those whose interpretation is dependent on a judge index or argument, we are faced with the question of why some of them – but not others – can also be interpreted as making factual statements, i.e. statements that can be evaluated as objectively true or false. In fact, it is not clear how they can acquire factual interpretations at all.

A related and deeper issue relates to the assumption underlying the above accounts that what we have called subjectivity relates inherently to the diverging perspectives of

distinct speakers. For classic personal taste predicates such as *tasty* and *fun*, this seems at least intuitively correct. But when we expand our focus to the full range of adjectives considered in the present work, it does not so clearly hold; rather, it seems that a single speaker’s judgments with respect to orderings may be potentially uncertain or variable. Consider for example two shirts, one which is clean except for a grass stain on the sleeve, the other slightly dingy overall. Which one do I consider dirtier, and which cleaner? I think my answer has to be ‘it depends’ – on what type of shirt and how it will be used, on what sort of dirt we are most concerned about, and so forth. The same might be said, for example, regarding which of two surfaces is rougher, or which of two fences is straighter. Formally, relativization to a judge is thus neither necessary or sufficient to account for the potential of variable judgments regarding orderings.

In recent work, attention has been drawn to the fact that adjectival subjectivity comes in more than one variety (see especially Sæbø 2009; Kennedy 2013; Bylina 2014; McNally and Stojanovic 2015). Most basically, it has been observed that while a wide variety of gradable adjectives are subjective in their positive forms (cf. (1)), only a subset of these are subjective in the comparative (cf. (2)). The conclusion is that there are two distinct loci for subjectivity. For vague gradable adjectives such as *tall*, subjectivity is localized in the semantics of the positive morpheme *pos*, and not the adjective itself. In adjectives such as *tasty*, *fun*, it derives from the lexical semantics of the adjective.

Kennedy proposes that this difference in which adjectival forms can be interpreted subjectively corresponds more fundamentally to two distinct types of subjectivity, the first deriving from uncertainty in the determination of the contextual standard for the application of a vague adjective, the second deriving from what he terms the “shared semantics of qualitative assessment”. This difference is reflected not only in presence or absence of ordering subjectivity but also in the possibility of embedding under *find*, which is possible for lexically subjective adjectives, but for at least some speakers infelicitous with the positive forms of vague gradable adjectives (e.g. *I find the chili tasty* vs. *?I find Anna tall*). He notes however that the two sorts of subjectivity might nonetheless be unified as deriving from a more basic property of ‘dimensional uncertainty’. For adjectives of the *tall* class, it is uncertainty as to the dimensions involved in standard calculation, while for those of the *tasty* sort, it is uncertainty as to how the dimensions of qualitative assessment are integrated by different judges.

Kennedy makes the further important observation that many gradable adjectives are ambiguous between an objective/dimensional reading and a subjective/qualitative reading. For example, to say that the cake is heavy might be to say something about its objectively measurable weight, or alternately about the subjective experience of tasting it. This suggests an account of the ‘mixed’ group found in the present experiment in terms of ambiguity (though we will see below that it is somewhat more complicated than this).

The notion of multidimensionality as a source of subjectivity is taken up further by Bylina (2014) and McNally and Stojanovic (2015), who observe that the class of adjectives exhibiting ordering subjectivity can itself be further subdivided. Bylina notes in particular that subjective readings for the comparative are possible for both adjectives such as *fun* and *interesting* that refer to internalized experiences as well those such as *intelligent* that do not; but only the former allow a judge or experiencer PP:

- (15) a. The chili was tasty to me.
 b. The book was interesting to/for me.
 c. ?? Anna is intelligent to/for me.

Bylinina proposes that the interpretation of both sorts of adjectives is dependent on a judge index, but that the judge plays a different role in the two cases. Members of the *tasty* class have an experiencer argument that is equated to the judge. In the case of adjectives such as *intelligent*, she draws on work by Sassoon (to be discussed further below) in proposing that their subjectivity derives from multidimensionality: degrees of intelligence, for example, can be conceptualized as the lengths of vectors in a multidimensional space, with the weights assigned to component dimensions being relativized to judges:

$$(16) \quad \llbracket m_x, \mathbf{intelligent} \rrbracket^{c:w,t,j} = \lambda.x \sqrt{\sum_Q [w_Q^j (m_{x,Q} \succ s_Q)]^2}$$

McNally & Stojanovic provide evidence for a further distinction between adjectives with experiential semantics such as *interesting*, *fun* and *tasty* and those such as *good*, *bad* and *beautiful* that are evaluative in the sense of expressing an attitude of positive or negative evaluation on the part of the speaker. It is the former group that are most felicitous with judge PPs, while the latter are marginal (here they cite the infrequency of *beautiful to me* outside of song titles); and while both sorts embed under *find*, in the case of the aesthetic adjective *beautiful*, this implies the evaluation is based on a specific perceptual experience, something that otherwise is not necessarily the case. Both factors however – experiential semantics and evaluativity – result in subjectivity.

Taking this work as a whole, we see that there are multiple sources that have been proposed for ordering subjectivity. In what follows I will seek to use the results of the experiment to shed light on the relative role of each, and the relationship between them. First, though, we will take a closer look at what we mean by multidimensionality.

4 Identifying multidimensionality

In this section, we will delve more deeply into the issue of multidimensionality in gradable adjectives. If we are to investigate the hypothesis that a source of subjectivity (including ordering subjectivity) is the multidimensional nature of the predicates in question, then we must have a way of identifying which adjectives are multidimensional. We will see that this is less straightforward than it might initially seem.

The most in-depth investigation of multidimensionality is found in the work of Sassoon (2007, 2011, 2012, 2013, 2015), who develops a comprehensive semantic theory that encompasses both multidimensional adjectives and nouns, and that extends to topics including the nature of the adjectival antonymy relationship and the semantics of comparison and degree modification. In Sassoon’s theory, multidimensional adjectives such as *healthy*, *sick*, *identical*, and *intelligent* are associated with dimensions that can be specified overtly or bound by explicit or implicit logical binding operators. For conjunctive adjectives such as *healthy*, the default binding operator is universal quantification: to be healthy is to be healthy in all contextually relevant respects (17a). For disjunctive adjectives such as *dirty*, the default is existential quantification: to be sick is to be sick

in some relevant respect (17b). Adjectives such as *intelligent* are mixed, with pragmatics determining the binding operation.

- (17) a. *healthy*: $\lambda x.\forall Q \in DIM(\textit{healthy}) : Q(x)$
 b. *sick*: $\lambda x.\exists Q \in DIM(\textit{sick}) : Q(x)$

Comparatives can then be analyzed as involving the counting of or quantification over dimensions; one individual might be evaluated as healthier than another if she is healthy in a larger number of relevant respects, if for relevant respects generally she is healthier, or if she is healthier in some contextually salient respect (Sassoon 2015).

Multidimensionality manifests itself grammatically in a number of ways: individual dimensions may be specified via prepositional phrases headed by *with respect to* or *in* (18) or inquired about via a *wh*-phrase (19); dimensions may be quantified over (20); and quantificational force can be restricted by exception phrases (21).³ None of these are possible with (one-)dimensional adjectives such as *tall*.

- (18) a. The patient is healthy with respect to blood pressure.
 b. The boxes are identical in size and weight.
 c. # Zoe is tall with respect to height.
- (19) a. In what respects is the patient healthy/sick?
 b. In what respects are the boxes identical?
 c. #? In what respect is Zoe tall?
- (20) a. The patient is healthy in every/most/three/some (important) respect(s).
 b. The boxes are identical in every/most/three/some respect(s).
 c. # Zoe is tall in every/most/?three/some respect(s).
- (21) a. The patient is healthy/not sick except for high blood pressure/asthma/a slight cold.
 b. The boxes are identical except for size/color.
 c. # Zoe is tall except for ...

Sassoon backs up these judgements with extensive corpus and experimental data, particularly relating to the pattern in (21).

Among the multidimensional adjectives that Sassoon investigates are a number that were found in the present research to exhibit ordering subjectivity: *good*, *bad*, *beautiful*, *ugly*, *happy*, *intelligent*, *tasty*, *clean* and *dirty*. Furthermore, when we look more generally at the mixed and especially the purely subjective groups that emerged from the experiment, we see that many or even most are multidimensional at least in a conceptual sense. Whether an individual or experience might be characterized as fun, interesting, boring, or easy – or more fun/interesting/boring/easy than another – is clearly dependent on multiple aspects or properties of the entities under consideration. Even the adjective

³Which quantifiers are felicitous, and whether an exception phrase is possible with an adjective in its positive or negated form, depend to some extent on whether the adjective is conjunctive or disjunctive. I will attempt as much as possible to abstract away from this issue here.

salty might be put in this class: while one might think that degree of saltiness is dependent on a single factor, namely salt content, research in psychophysics has in fact found that perceptions of saltiness are impacted by a variety of other factors, including consistency, texture and fat content (see e.g. Christensen 1980; Pflaum et al. 2013; Suzuki et al. 2014).

However, when we attempt to confirm multidimensional status of such adjectives via tests based on the constructions in (18)-(21), and thereby clarify which of the adjectives exhibiting ordering subjectivity are multidimensional, the results are inconclusive at best. This point is seen particularly clearly with the personal taste predicates *tasty* and *fun*. Consider the examples in (22) and (23):

- (22) a. The chili was tasty with respect to ...
 b. In what respect/way was the chili tasty?
 c. The chili was tasty in every/?most/??three/some respect(s).
 d. The chili was tasty except for the consistency/being too salty/??
- (23) a. The roller coaster was fun with respect to ...
 b. In what respect was the roller coaster fun?
 c. The roller coaster was fun in ?every/?most/??three/some respect(s).
 d. The roller coaster was fun except for the wind/the rattling/??

It is hard to imagine how the sentences in (22a),(23a) might be felicitously continued, or the questions in (22b), (23b) felicitously answered. What are the respects of tastiness and fun that contribute to the attribution of these predicates? At most, the questions seem to allow a rhetorical interpretation, challenging the listener to name even one ground for calling the chili tasty or the roller coaster fun. Similarly, while universal and existential quantification over dimensions seem at least moderately acceptable ((22c),(23c)), producing emphatic and hedging effects, respectively, precise counting of dimensions (*??fun/tasty in three respects*) is decidedly odd. Finally, while it may be possible to distinguish a few particular aspects of the properties in questions to form the basis of exception phrases (e.g. saltiness and consistency in the case of *tasty*), after these the task becomes more difficult (see (22d),(23d)), suggesting that there is a considerable residual meaning that cannot be easily divided into discrete sub-dimensions.⁴ The conclusion would seem to be that for some intuitively multidimensional predicates, their individual component dimensions are less grammatically – or even conceptually – accessible than those for adjectives such as *healthy* and *identical*.

The same issue emerges with other sorts of evaluative predicates, where we see that even when examples parallel to (18)-(21) sound felicitous, they do not necessarily involve specification of or quantification over dimensions. Take for example *beautiful*, one of the adjectives that fell in the purely subjective group in our experiment. A Google search yields thousands of examples of the phrases *beautiful in every respect* and *beautiful in every way*. But many of these have the character of those in (42), where the listed aspects are not component dimensions of the predicate *beautiful* but rather component parts of a complex entity or event that is the subject of predication.

⁴Consistent with this observation, in Sassoon (2012), *tasty* is found experimentally to be less acceptable with exception phrases than many of the other multidimensional adjectives tested.

- (24) a. Nicola & Marc’s wedding was beautiful in every respect ... the weather, the dress, the venue, the cars and most of all ... the people.
 b. This newly constructed home is beautiful in every way, featuring a great kitchen, an unbelievable screened porch and generous living space.

Something similar is seen with exception phrases: *Zoe is beautiful except for ...* is most naturally continued with something like *her crooked nose / her small eyes / her hair / etc.*; but nose, eyes, hair and the like are not dimensions of beauty but rather parts of the individual described. To be sure, dimensional uses can be found, as when we characterize a painting as *beautiful except for the color* (McNally and Stojanovic 2015). But the simpler the object of predication, the more difficult it is to construct such examples. As an extreme case, imagine a paint chip in a particular shade of blue. I might characterize the color as beautiful, but it is hard to imagine specifying the dimensions that make it so (*?this color is beautiful with respect to...*) or less so (*?this color is beautiful except for...*); to ask in which respect it is beautiful would be odd. Replacing *beautiful* with *ugly* makes these judgments in my opinion even sharper. Sassoon (2013) acknowledges and discusses non-dimensional uses of exception phrases with multidimensional adjectives, but without, I think, fully recognizing difficult of creating true dimensional examples for those such as *beautiful*.

Consider now adjectives in our mixed group. Of these, *clean* and *dirty* are discussed as multidimensional by Sassoon, and this seems intuitively correct: it is relatively easy to imagine the sort of dimensions that contribute to determining an entity’s degree of dirtiness or cleanness; for a shirt, for example, these might be amounts of different sorts of dirt, soil, odor and the like. Correspondingly, examples such as the following demonstrate that *clean/dirty* largely pass the above-described tests for multidimensionality:

- (25) a. In what respect(s) was the shirt clean/dirty?
 b. The shirt was clean/dirty every/most/??three/some respect(s).
 c. The shirt was clean / wasn’t dirty except for the musty smell / a few grass stains / being slightly dingy.

But when we look at other members of this group, we again find that these tests do not serve us as well. Taking *except* phrases as an example, it is difficult to construct true dimensional completions of examples such as the following:

- (26) a. The line was(n’t) straight/curved except for ...
 b. The leather was(n’t) smooth/rough except for ...
 c. The knife was(n’t) sharp/dull except for ...
 d. The soup was(n’t) salty except for ...

But there is nonetheless a sense in which adjectives such as these are multidimensional. This is most clearly brought out by considering cases of potential disagreement. For example, we might disagree – or simply find it difficult to decide – which of the two lines below is straighter or more curved, the issue being how exactly we should measure degree of straightness or curvature: is it a matter of the number of curves? the sharpness of each? the total area of deviation from perfect straightness? There seems to be no principled correct answer.



To take a more concrete example, imagine two city streets, one paved and completely smooth except for a few largish speed bumps and potholes, the other with a cobblestone surface. Which is bumpier? Again the answer seems to be ‘it depends’, the issue once more being how different sorts of bumps, dips and other deviations from complete flatness should be integrated to derive an overall degree of bumpiness.⁵ I believe similar examples might be constructed for other members of the mixed class, including *rough/smooth*, *sharp/dull* and perhaps even *wet/dry*. This does not seem to be multidimensionality in quite the same sense as that characterizing adjectives such as *healthy*, whose meanings can readily be broken down into discrete independent dimensions (e.g. blood pressure, cholesterol, etc.) that we can name, count and/or quantify over. But adjectives of the *curved* and *bumpy* type share with those of the *healthy* type the property that their attribution depends on multiple aspects of the physical characteristics of entities, which must be integrated in some way to produce the overall meaning of the adjective.

We have seen that there are adjectives that are in some sense multidimensional but that do not readily occur in the constructions in (18)-(21). The reverse is also true: certain adjectives that are generally considered to be dimensionally ambiguous rather multidimensional are relatively acceptable with *respect*. Examples are *large* and *long*:

- (28) In which respect is London larger than New York?
Land area ✓ Population size ✗
- (29) The sofa is larger than the bench in every respect.
- (30) a. The trip to Tübingen is longer than the trip to Konstanz.
 b. In which respect – travel time or distance in kilometers?

To summarize, the constructions exemplified in (18)-(21) provide evidence for the existence of grammatically relevant distinctions between dimensions such as height and health, which must be accounted for in any comprehensive theory of scalar meaning. But importantly, tests of this sort do not pick out a single uniform class of multidimensional adjectives. Instead, when we look beyond clear cases of unidimensional adjectives such as *tall*, what we find is a spectrum, ranging from those adjectives whose dimensions can be readily identified and referenced to those whose component dimensions are much less grammatically or even conceptually accessible. Furthermore, the tests themselves differ in the results they produce. Existential or universal quantification over dimensions (*in every/some respect*) is possible with a wide variety of adjectives that we might tend to classify as multidimensional (and even some that we do not); constructions that require us to name individual dimensions (*in what respect?*; *with respect to ...*; *except for ...*) and especially count dimensions (*in three respects*) are compatible with progressively more restricted subclasses.

Are we left then to conclude that the only guide to classifying an adjective as multidimensional is our intuitions regarding the number of aspects or factors involved in its

⁵The pair *flat/bumpy* was not included in the present experiment, but I hypothesize that they would behave similarly to pairs such as *smooth/rough*; as *bumpy* provides a particularly nice example, I allow myself the liberty of using it here.

ascription, and further that the class so identified is anything but homogenous? Perhaps to some extent yes, but I would like to suggest that there are nonetheless two further distinctions that correlate with those discussed above, and that are relevant to the proper semantic analysis of such adjectives.

The first of these relates to measurability. For classic examples of multidimensional adjectives such as *healthy* and *identical* as well as members of our mixed group such as *clean/dirty*, *straight/curved* and *flat/bumpy*, the different component dimensions or aspects of an entity that contribute to the adjective's meaning can – at least in principle – be numerically measured: we can, for example measure the health of an individual's cardiovascular system, the size and weight of a box, the amount of different sorts of dirt on a shirt, the number and degree of curves in a line, and so forth. For predicates of personal taste such as *fun* and *tasty* and evaluative predicates such as *beauty*, this is less clearly the case. We might for example have the intuition that draftsmanship is one of the factors that contribute to evaluating a painting as beautiful, but draftsmanship itself seems as immune to numerical measurement as beauty itself is. Correspondingly, for adjectives of the first sort we can conceive of their dimensions being integrated by mathematical operations such as summation or averaging; for those of the second sort, the manner in which their component dimensions might be integrated is far less transparent.

A second and more fundamental difference involves the relationship between the component dimensions and the overall meaning of the adjective. This is again most easily seen in examples featuring the comparative. Consider first the examples in (31). If a speaker asserts one of these statements, I believe it is reasonable for her interlocutor to ask for specifics via a *respect* question.

- (31) a. Fred is healthier/sicker than Tom.
 b. The red shirt is cleaner/dirtier than the blue one.
 c. Weserstrasse is bumpier than Friedelstrasse.
 i. In what respect / way?

But for assertions based on the comparative forms of evaluative adjectives and personal taste predicates, such a question about respects is, as I have suggested above, slightly infelicitous. Instead, the most natural way to question the speaker's assertion is to ask for her reasons for it, for example with *What makes you say that?*

- (32) a. The chili is tastier than the soup.
 b. The roller coaster was more fun than the ferris wheel.
 c. The Picasso is more beautiful than the Miró.
 i. #In what respect / way?
 ii. Why do you say so / what makes you say that?

These judgments are subtle, but nonetheless I think quite real, and they point to a difference in the role played by the dimensions that correspond to the two sorts of adjectives. The meaning of adjectives such as *healthy*, *clean/dirty* and *flat/bumpy* is in a sense built up from their component dimensions, integrated in some appropriate way. But for adjectives such as *fun*, *tasty* and *beautiful*, what we have called dimensions are more properly factors that contribute to an agent's subjective experience with or evaluation of an entity or event. That is, the adjective's meaning is not a direct function of its dimensions;

rather, ‘dimensions’ serve as the basis for a taste, value or aesthetic judgment, and it is this that might more properly be considered the meaning of the adjective.

This latter point is similar to one made by the moral philosopher Hare (1952) in discussing the meaning of *good* (see also Umbach to appear for related discussion). Hare argues that evaluative terms such as *good* have the special function in language of commending, and cannot be defined in terms of other words which themselves do not have this function without losing the means of performing the commending function. A good strawberry, for example, may be one that is large, red and juicy; but *good* as applied to strawberries cannot be defined as meaning ‘large, red and juicy’, because in that case the sentence *A strawberry that is large, red and juicy is a good strawberry* would be a tautology. Hare further argues for the need to distinguish the meaning of evaluative words from the criteria for their application; the latter vary with the class of items to which the word is applied (i.e. what makes a good car is different from what makes a good strawberry), while the meaning, whose core is the commending function, remains constant. Criteria as discussed by Hare are close in spirit to what we have called the dimensions of evaluative adjectives.

To return to the semantics of multidimensional adjectives, we have seen that those such as *healthy* and *sick* can be represented in terms of their component dimensions, as in (14), and it is reasonable to extend this also to members of our mixed class such as *clean/dirty*. But if we are to take the above discussion seriously, it is less clear that this should be done for adjectives such as *beautiful* and *fun*, where dimensions are not direct components of the adjective’s meaning but rather criteria that form the basis of an evaluative or taste judgement. I will assume this distinction in what follows.⁶

5 Proposal

In this section, I will outline a theory of gradable adjective meaning that formalizes the observations from the previous sections, and that provides the basis for explaining the availability of subjective and objective readings of the comparative forms of different sorts of adjectives.

5.1 Scalar semantics

I begin with the definition of a scale S as triple of the following form:

- (33) $S = \langle D, \succ, DIM \rangle$, where
- DIM is a dimension of measurement
 - D is a set of degrees
 - \succ is an ordering relation on D

Importantly, this conceptualization establishes a one-to-many rather than one-to-one relation between dimensions of measurement and scales, such that a single dimension DIM

⁶In this, I differ from Umbach (to appear), who implements Hare’s notion of criteria via multidimensional attribute spaces, thereby aligning the interpretation of evaluative adjectives to that of dimensional adjectives.

may be tracked by multiple scales that differ in their structures. Differing from some other authors, I assume here that D can but need not be the real numbers, and that the ordering relation \succ can but need not be a total order on D .

A measure function μ_{DIM} can then be defined as a function from some domain of measurement Dom (e.g. the domain of individuals or of events, or some subset thereof) to a scale S tracking dimension DIM .

Building on proposals by Sassoon (2010) and Kennedy (2013), I then propose that gradable adjectives have underspecified semantics, lexicalizing not a single measure function but a family of functions indexed to contexts c . Each context specifies a world, time and judge as well as other aspects of the situation of utterance; here I explicit assume that two contexts c and c' may differ in the measures assigned to individuals, even if the physical properties of objects in the world remain the same. Thus the general template for gradable adjective meaning is the following:

$$(34) \quad \llbracket \text{Adj} \rrbracket^c = \lambda d \lambda x. \mu_{DIM}^c(x) \succeq d$$

To put this differently, gradable adjectives on this view lexicalize dimensions rather than particular scales. The presence or absence of ordering subjectivity can then be related to the further constraints on the set of measure functions μ_{DIM}^c encoded by the adjective.

5.2 Sources of objectivity

Above we noted the link between measurability – i.e., the possibility of associating entities with numerical measures – and objective rather than subjective interpretations for the comparative. Building on this insight, I propose that objective readings are possible in those cases where μ_{DIM}^c is such that it allows a principled, order-preserving mapping to the real numbers. This has the effect of externalizing orderings of individuals, aligning them across speakers to the fixed order of the number line.

Such a mapping can come about in several ways. The most straightforward of these arises when the adjective lexicalizes measure functions that are additive with respect to concatenation, meaning that the measure assigned to two individuals concatenated in the relevant way is the sum of their two individual measures. The dimension of height is a classic example: the height of two individuals stacked one on top of the other is the sum of their individual heights. Additivity may be encoded via a constraint on μ_{DIM}^c . We define a concatenation operation \oplus on elements of the domain, and a corresponding addition operation $+$ on degrees on the scale. The constraint is then that in (35); a sample denotation for an adjective satisfying this constraint is (36).

$$(35) \quad \textbf{Additivity measure functions:}$$

$$\forall a, b \in Dom, \mu_{DIM}^c(a \oplus b) = \mu_{DIM}^c(a) + \mu_{DIM}^c(b)$$

$$(36) \quad \llbracket \text{tall} \rrbracket^c = \lambda d \lambda x. \mu_{HEIGHT}^c(x) \succeq d,$$

$$\text{where } \forall a, b \in Dom, \mu_{HEIGHT}^c(a \oplus b) = \mu_{HEIGHT}^c(a) + \mu_{HEIGHT}^c(b)$$

Other dimensions that satisfy additivity include weight, depth, width, length, volume and duration. Even the dimension of cost arguably falls in this class: while items are often cheaper if purchased in quantity, the fact that we recognize this as a discount is an indication that we perceive cost as inherently additive.

Importantly, not all dimensions lexicalized by gradable adjectives are additive in the sense of height. Temperature is a prime example (Lassiter 2011): the temperature of two bowls of soup poured together is not the sum of their two individual temperatures, but rather somewhere intermediate between them. Among adjectives that lexicalize non-additive dimensions, there are nonetheless at least two other possible routes to numerical measurement.

First, there are dimensions for which natural, speaker-external phenomena serve as the basis for units of measure. Examples of this include temperature as well as temporal dimensions. In the case of time, the rotation of the earth and its orbit around the sun provide the basis for the units ‘day’ and ‘year’; subdivision and concatenation of these units yield further units such as ‘hour’, ‘minute’, and ‘week’. For temperature, the freezing and boiling points of water provide two anchor points on the scale, which can then be divided into equal increments, for instance by equal increases in the level of mercury in a thermometer. Units derived in this way provide another sort of principled mapping from entities to numbers.

A further class of dimensions that support numerical measurement consists of those whose corresponding measure functions can be built up from measure functions of the above two classes in a context-independent way. The dimension of fullness provides a good example: the degree of fullness of a container (say, a bottle or gas tank) can be expressed as the volume of its contents divided by its capacity, i.e., the volume it is able to hold. A half full tank, for example, is one whose contents have half the volume of its capacity.

Adjectives falling in this class are those that satisfy the constraint in (37); as an example, the corresponding lexical entry for the adjective *full* is given in (38):

(37) **Context independent derived measure functions:**

$$\mu_{DIM}^c(x) = f(\mu_{DIM_1}^c(x), \mu_{DIM_2}^c(x), \dots, \mu_{DIM_n}^c(x)),$$

where $\mu_{DIM_1}^c, \mu_{DIM_2}^c, \dots, \mu_{DIM_n}^c$ are objective measure functions

(38) $[[full]]^c = \lambda d \lambda x. \mu_{FULLNESS}^c(x) \succeq d,$

$$\text{where } \forall x \in Dom, \mu_{FULLNESS}^c(x) = \frac{\mu_{VOLUME}^c(content(x))}{\mu_{VOLUME}^c(capacity(x))}$$

As defined here, fullness is not strictly speaking additive: two half full glasses when placed together do not produce a full glass, though their combined contents are the same as that of a single full glass. But because degrees of fullness can be derived via a fixed function of two additive measure functions, a principled mapping to numbers can nonetheless be derived, allowing numerical/proportional measure expressions such as *20% full*. Other dimensions in this class might be purity (defined as volume of impurities relative to total volume) and speed (distance traveled divided by duration), and as expected these too allow numerical measures (*90% pure*, *5 kilometers per hour faster/slower*).

In all of these cases, entities can be associated in a principled way with numerical values that reflect their position with respect to the relevant dimension *DIM*. The prediction is that the comparative form of the corresponding adjectives will be interpreted objectively, and this is consistent with our experimental findings for *tall/short*, *expensive* (additive dimensions), *old/new* (time expressions) and *full/empty* (function of additive measure functions). We would predict similar results for other adjectives in these classes.

5.3 Sources of subjectivity

Let us turn now to the characteristics of adjectives whose comparative forms can be interpreted subjectively, as diagnosed by the possibility of faultless disagreement. Building on the previously discussed observations by Bylina (2014), as well as the discussion in Section 4, I propose there are two distinct sources of ordering subjectivity, namely multidimensionality and judge dependence.

Multidimensionality. Above we discussed the insight that certain adjectives exhibiting ordering subjectivity are multidimensional. Underspecification in or uncertainty about the component dimensions and how they should be integrated results in the potential for disagreement as to orderings. Take for example the pair *clean/dirty*. Intuitively, the degree of cleanness or dirtiness of an object is a function of the amount and type of dirt on it, perhaps in proportion to its size. But which sorts of dirt (broadly construed) we are concerned with, and how different sorts should be weighted relative to one another, are matters of potential disagreement, and there does not seem to be a principled correct choice. On one way of making this more specific, shirt *a* might work out to be dirtier than shirt *b*, while on another equally valid choice, the reverse relation might obtain.

To formalize this, I follow Sassoon (2013) and Bylina (2014) in proposing that adjectives of this sort are associated in each context c with a set of component dimensions $DIM_1^c, DIM_2^c, \dots, DIM_n^c$. Departing somewhat from these authors, I further assume that to each dimension DIM_i^c there corresponds a measure function $\mu_{DIM_i^c}^c$, the outputs of which are integrated by some function f^c . We have already seen something similar in the form of the lexical entry for *full*. But in that case, subjectivity did not arise, because both the component dimensions and the manner of their combination were fully specified. Ordering subjectivity arises when this requirement is relaxed, such that one or both of these factors becomes context dependent. (39) specifies the form of such functions, and (40) gives a plausible if undoubtedly overly simplistic entry for *dirty* in this form.

(39) **Context-dependent derived measure functions:**

$$\mu_{DIM}^c(x) = f^c(\mu_{DIM_1^c}^c(x), \mu_{DIM_2^c}^c(x), \dots, \mu_{DIM_n^c}^c(x))$$

(40) $[[\text{dirty}]]^c = \lambda d \lambda x. \mu_{DIRTINESS}^c(x) \succeq d,$

$$\text{where } \mu_{DIRTINESS}^c(x) = \frac{\sum_{i=1}^n k_i^c \cdot \mu_{AMOUNT}^c(\text{dirt}_i^c(x))}{\mu_{SIZE}^c(x)}$$

Note that the individual dimensions that underlie such entries may themselves be objectively measurable; subjectivity derives from the potential for variation in the choice of these dimensions and how they are combined.

Judge dependence. The entries in (39) and (40) do not explicitly reference a judge. Rather, measure functions are indexed to contexts; and distinct orderings in two contexts c and c' may derive from a difference between judges (the judge being part of the context), but also from other contextual factors. This is as it should be, given our previous observation that uncertainty or variability regarding the ordering of individuals relative to a multidimensional property such as dirtiness can persist in the judgments of a single speaker. However, we have also seen that many gradable adjectives denote properties whose ascription depends necessarily on what McNally and Stojanovic (2015) refer to as the mediation of a sentient individual. These include value judgments

(*good/bad*), aesthetic judgments (*beautiful/ugly*), taste judgments (*tasty*), experiential properties (*interesting/boring*) and internal states (*happy/sad*). Such adjectives do not directly describe properties of objects and events in the world, but rather our perceptions of, judgments about and experience with the objective world. For this class, I propose that their dependence on sentient mediation be represented in their semantics. I thus follow the existing tradition of work on subjectivity in taking these to involve measure functions parameterized to a judge. Adapting for concreteness the relativist approach, we may represent this as follows:

(41) **Judge dependent measure functions:**

$$\llbracket \text{Adj} \rrbracket^{c:j} = \lambda d \lambda x. \mu_{DIM}^{c:j}(x) \succeq d$$

where $\mu_{DIM}^{c:j}(x)$ should be interpreted as

‘the degree to which j judges x in context c to have property DIM ’

(42) $\llbracket \text{beautiful} \rrbracket^{c:j} = \lambda d \lambda x. \mu_{BEAUTY}^{c:j}(x) \succeq d$

Here I do not attempt to resolve whether this class can be further subdivided, and in particular whether we must distinguish between adjectives with experiential semantics such as *tasty* and *interesting* and evaluative predicates such as *beautiful* (per McNally and Stojanovic 2015, cf. Section 3). Further research in this area would be beneficial.

An important question to address at this point is whether we really need both of the possibilities provided in (39) and (41), or whether these could be collapsed together into a single more general template for subjective gradable adjective meaning that includes context, judge and component dimensions. This is a somewhat difficult question to resolve empirically. In Section 4 we saw the issues involved in finding adequate tests for multidimensionality. Likewise, plausible tests for the presence of a judge parameter or argument, in particularly the possibility of embedding under subjective attitude verbs and modification by a judge PP, might in fact pick out only a subset of this class, namely those involving an experiencer as part of their semantics (see Section 3). But with these caveats in mind, there are nonetheless both empirical and formal reasons to maintain a distinction along the lines of that between (39) and (41).

The first part of the question is whether adjectives whose interpretation is explicitly dependent on a judge or experiencer should also be represented as explicitly multidimensional, contra (41). The discussion at the end of Section 4 suggests that this should not be done. The ‘dimensions’ underlying adjectives such as *beautiful* and *tasty* are first of all not as accessible grammatically as for paradigm cases such as *healthy*, and are furthermore themselves largely non-measurable, making it questionable that they could be integrated mathematically as in (39). More fundamentally, I argued above that dimensions play a different role for such adjectives than for those such as *healthy* and *clean/dirty*, being not components of the adjective’s meaning but rather grounds for an agent’s taste, value or aesthetic judgment; this suggests that in these cases the meaning of the adjective should not be represented as a function of its dimensions. On the basis of these observations, as well as general considerations of parsimony, I thus conclude that subjective adjectives of the judge-dependent type are only multidimensional at the conceptual level, but should be represented as unidimensional in their semantics.

The flip side of the question is whether multidimensional adjectives should also be represented as explicitly judge dependent. Evidence against this comes from the existence

of adjectives that are ambiguous between the two types of subjective interpretations. Consider again the adjective *bumpy*, and the two city streets from our earlier example, one flat except for isolated potholes and speed bumps, the other with a cobblestone surface. On the basis of this description alone (or perhaps pictures of the two streets), there is room for uncertainty or between-speaker disagreement as to which of the sentences in (43) is true, the issue being how to weight the different sorts of bumps to arrive at an overall measure of bumpiness. But the disagreement in (44) implies something further, namely that the speakers have experienced the two streets in question (e.g. by riding a bike over them):

- (43) a. Weserstrasse is bumpier than Friedelstrasse.
 b. Friedelstrasse is bumpier than Weserstrasse.
- (44) a. I find Weserstrasse bumpier than Friedelstrasse.
 b. I find Friedelstrasse the bumpier of the two.

Or for another example, on examining two knives under a microscope, one might be uncertain as to which is sharper, the issue again being what counts as sharper. But to assert that *I find the first knife sharper than the second one* requires that I have used both of them, and says something about my own subjective and experience-based perception of how the two should be ordered. The difference between the examples with and without *find* suggests that an adjective can be multidimensional – and thus exhibit ordering subjectivity – without explicitly having a judge or experiencer as part of its semantics.

5.4 Mixed predicates

Having discussed the characteristics of the measure functions that support objective and subjective judgments about orderings, let us now relate these further to our experimental results, and in particular to the finding that a large class of gradable adjectives allow both sorts of interpretations for their comparative forms.

One obvious possibility is that adjectives of this sort are ambiguous between the two sorts of interpretations, as suggested by Kennedy (2013). This would not require us to posit a lexical ambiguity for the adjectives in question; rather, it is a possibility already allowed for by the underspecified, context-dependent template for gradable adjective meaning in (34). An explanation in terms of ambiguity is plausible in particular for adjectives whose meaning relates to perceptual dimensions such as sight, hearing and taste. A good example is *salty*, which might be interpreted objectively in terms of salt content or subjectively in terms of an experiencer’s perception of a substance’s taste properties; on the former interpretation *salty* aligns to adjectives of the *full* class (allowing numerical measures), while on the latter it aligns to the judge or experiencer-dependent *tasty/beautiful* class. Correspondingly we found mixed judgments in the experiment. Other adjectives tested that might fall in this group include *light/dark* and *hard/soft*.

However, I believe this is not the only possibility; rather, the logical forms of adjectives of the multidimensional sort themselves allow objective as well as subjective interpretations. The intuition behind the formalisms in (39) and (40) is that the lexical entries of adjectives such as *clean/dirty* underspecify the component dimensions and their manner

of combination that go into the assessment of an entity’s degree of (say) cleanness or dirtiness. Ordering subjectivity arises when two speakers disagree about how these should be specified (i.e., assume different contexts), or when a single speaker is uncertain as to how to specify them (i.e., is uncertain about the precise context). But contexts can only vary so much: while there may be room for disagreement as to how different sorts of dirt and so forth should be weighted to arrive at an overall degree of dirtiness for an entity, a shirt that is covered with oil stains must be evaluated as dirtier than one that is clean except for a few smudges of dirt near the hem. In formal terms, for all contexts c , the order of the degrees assigned to these two shirts relative to the dimension dirtiness remains the same. One possibility thus is that that subjects who gave ‘fact’ judgments for adjectives such as these assumed entities that were different to such a degree. A second possibility for the objective interpretation of a multidimensional adjective is that speakers assume the context is such that the component dimensions and their manner of interpretation is fully determined. Such might be the case in a laboratory testing situation, where we have some agreed on way of quantitatively measuring degree of cleanness/dirtiness. Further research is needed to better understand the role of these two factors in producing the mixed results in our experiment.

Note that for judge-dependent adjectives of the *beautiful* sort, objective interpretations for the comparative cannot be derived in the same way. For members of the multidimensional class, the range of possible variation in degrees assigned to entities is constrained by the possible choices for the component dimensions $DIM_1^c, DIM_2^c, \dots, DIM_n^c$ and the function f^c ; for certain pairs of entities or situations of utterance, these constraints have the effect that the possibility of ordering subjectivity disappears. But for *beautiful* and the like, varying judgments about orderings derive directly from the varying perceptions and tastes of distinct agents or experiencers, which are not constrained in any formal way. Correspondingly, we predict members of this class to be interpreted purely subjectively in the comparative, and this is exactly what we found.

In Table 1, the observations from this section are summarized in the form of a classification of gradable adjectives by the formal properties of their (families of) measure functions, and the corresponding availability of objective versus subjective readings for the comparative. The table is populated with examples taken from the above discussion. I have not, however, attempted a full classification of all the adjectives experimentally tested, and here there are questions that could be raised; as an example, on the criteria discussed above the adjective *intelligent* would seem to be multidimensional rather than judge-dependent, but unlike others in this class it elicited purely subjective readings for its comparative form. Additional research would be beneficial to understanding if the categorization proposed here must be refined, and in further developing tools to assign adjectives to categories.

6 Conclusions

The starting point for this paper was the observation that the comparative forms of certain gradable adjectives are interpreted subjectively, a pattern that is problematic for standard theories of gradability. The hypothesis was explored that subjectivity of this sort derives from the multidimensional nature of the properties in question. I have

Table 1: Classes of gradable adjectives

		Interpretation of Comparative	
		Objective	Subjective
(a) Measurable		✓	
- Additive	<i>tall/short, expensive</i>		
- Externally anchored	<i>new/old, hot/cold</i>		
- Context-independent derived	<i>full/empty, pure/impure, fast/slow</i>		
(b) Multidimensional (context-dependent derived)	<i>clean/dirty, straight/curved</i>	✓	✓
(c) Judge-dependent	<i>tasty, fun, beautiful/ugly</i> <i>interesting/boring, happy/sad</i>		✓
Ambiguous between (b) & (c)	<i>bumpy/flat, sharp/dull</i>	✓	✓
Ambiguous between (a) & (c)	<i>salty, hard/soft, dark/light</i>	✓	✓

attempted to make two empirical contributions in this work. The first is to demonstrate experimentally that ordering subjectivity is more widespread than previously recognized, and furthermore that adjectives with this property pattern into two groups, depending on whether or not they also allow objective readings for their comparative forms. The second is to show that multidimensionality is a complex and multifaceted phenomenon, and that not all gradable adjectives that are conceptually multidimensional should be represented as explicitly multidimensional in the semantics. From a theoretical perspective, I have argued that the facts are best captured by positing two distinct sources of ordering subjectivity, multidimensionality being one, the second being parameterization to a judge. Formally, I have expressed this insight in a theory of gradability on which the availability of objective vs. subjective readings of the comparatives derives from the formal properties of the measure functions lexicalized by gradable adjectives.

I have not attempted to apply this analysis to the positive form of subjective gradable adjectives, or to relate the discussion of ordering subjectivity to the large body of insights on subjectivity more generally. Clearly there are connections here, but attempting to explore them would take us too far from the central topic of the paper. I therefore leave the broader implications of these findings for the study of subjectivity as a topic for the future.

Acknowledgments

For helpful discussions I would like to thank Louise McNally, Galit Sassoon, Carla Umbach and other members of the audiences at ExpGrad, SuB20 and the ZAS. Thanks also to Nadja Reinhold and Lisa Reimann for assistance with experiment. All errors are of course my own. This work was supported by the Deutsche Forschungsgemeinschaft under grant 1157/1-1.

References

- Bates, Douglas M., Martin Mächler, Ben Bolker, and Steve Walker. 2014. Lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7. <http://CRAN.R-project.org/package=lme4>.
- Bylinina, Lisa. 2014. The grammar of standards: Judge-dependence, purpose-relativity, and comparison classes in degree constructions. Doctoral Dissertation, Utrecht University.
- Christensen, Carol M. 1980. Effects of solution viscosity on perceived saltiness and sweetness. *Perception and Psychophysics* 28:347–353.
- Crosswell, Max J. 1977. The semantics of degree. In *Montague grammar*, ed. Barbara H. Partee, 261–292. New York: Academic Press.
- Hare, R. M. 1952. *The language of morals*. Oxford: Clarendon Press.
- Heim, Irene. 2000. Degree operators and scope. In *Proceedings of the 10th Semantics and Linguistic Theory Conference*, ed. Brendan Jackson and Tanya Matthews, 40–64. Ithaca, NY: CLC Publications.
- Hothorn, Torsten, Frank Bretz, and Peter Westfal. 2008. Simultaneous inference in general parametric models. *Biometrical Journal* 50:346–363.
- Kennedy, Christopher. 1997. Projecting the adjective: The syntax and semantics of gradability and comparison. Doctoral Dissertation, University of California at Santa Cruz.
- Kennedy, Christopher. 2013. Two sources of subjectivity: Qualitative assessment and dimensional uncertainty. *Inquiry: An Interdisciplinary Journal of Philosophy* 56:258–277.
- Kennedy, Christopher, and Louise McNally. 2005. Scale structure, degree modification and the semantics of gradable predicates. *Language* 81:345–381.
- Lasersohn, Peter. 2005. Context dependence, disagreement, and predicates of personal taste. *Linguistics and Philosophy* 28:643–686.
- Lassiter, Dan. 2011. Measurement and modality: the scalar basis of modal semantics. Doctoral Dissertation, New York University.
- McNally, Louise, and Isidora Stojanovic. 2015. Aesthetic adjectives. In *The semantics of aesthetic judgment*, ed. James Young. Oxford: Oxford University Press.
- Pflaum, Tabea, Katharina Konitzer, Thomas Hofmann, and Peter Koehler. 2013. Influence of texture on the perception of saltiness in wheat bread. *Journal of Agricultural and Food Chemistry* 61:10649–10658.
- R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.

- Sæbø, Kjell Johan. 2009. Judgment ascriptions. *Linguistics and Philosophy* 32:327–352.
- Sassoon, Galit. 2007. Vagueness, gradability and typicality, a comprehensive semantic analysis. Doctoral Dissertation, Tel Aviv University.
- Sassoon, Galit. 2010. Measurement theory in linguistics. *Synthese* 174:151–180.
- Sassoon, Galit. 2011. Adjectival vs. nominal categorization processes: The rule vs. similarity hypothesis. *Belgium Journal of Linguistics* 25:104–147.
- Sassoon, Galit. 2012. The double nature of negative antonymy. In *Proceedings of Sinn und Bedeutung 16. MIT Working Papers in Linguistics*, ed. Ana Aguilar Guevara, Anna Chernilovskaya, and Rick Nouwen, 543–556.
- Sassoon, Galit. 2013. A typology of multidimensional adjectives. *Journal of Semantics* 30:335–380.
- Sassoon, Galit. 2015. A degree-approach account of multidimensional gradability. Unpublished manuscript.
- Schwarzschild, Roger. 2005. Measure phrases as modifiers of adjectives. *Recherches linguistiques de Vincennes* 34:207–228.
- Solt, Stephanie. 2016. Ordering subjectivity and the absolute/relative distinction. In *Proceedings of Sinn und Bedeutung 20*.
- Stephenson, Tamina. 2007. Judge dependence, epistemic modals, and predicates of personal taste. *Linguistics and Philosophy* 30:487–525.
- Stojanovic, Isidora. 2007. Talking about taste: Disagreement, implicature arguments, and relative truth. *Linguistics and Philosophy* 30:691–706.
- Suzuki, A.H., H. Zhong, J. Lee, and S. Martini. 2014. Effect of lipid content on saltiness perception: A psychophysical study. *Journal of Sensory Studies* 29:404–412.
- Umbach, Carla. to appear. Evaluative propositions and subjective judgments. In *Subjective meaning*, ed. J. van WijnbergenHuitink and C. Meier. Berlin: de Gruyter.