Multidimensionality, Subjectivity and Scales: Experimental Evidence

Stephanie Solt

Abstract This paper investigates the subjective interpretation of the comparative forms 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 of an experimental study are presented which demonstrate 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. Evidence is also offered that adjectival multidimensionality itself is a heterogenous phenomenon. 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. This proposal is formalized within a semantic framework in which gradable adjectives lexicalize families of measure functions indexed to contexts and in some cases judges.

Keywords Adjective · Comparative · Gradability · Measurement · Multidimensionality · Subjectivity · Faultless disagreement · Predicate of personal taste · Context dependence

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 (see Kölbel 2004; Lasersohn 2005, 2009; Stephenson 2007; Sæbø 2009; Moltmann 2010; and other work cited below). Such 'faultless disagreement' is observed most classically with so-called predicates of personal taste such as *tasty* and *fun*, but also with evaluative adjectives more generally (e.g. *beautiful*) and with the unmodified positive forms of vague gradable adjectives (e.g. *tall*):

S. Solt (✉)
Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS), Berlin, Germany
e-mail: solt@leibniz-zas.de

E. Castroviejo et al. (eds.), *The Semantics of Gradability, Vagueness, and Scale Structure*, Language, Cognition, and Mind 4, https://doi.org/10.1007/978-3-319-77791-7_3

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

Recently, attention has turned to a second sort of subjectivity, which characterizes the comparative forms of some but not all gradable adjectives (Kennedy 2013; Bylinina 2014, 2017; Umbach 2016; McNally and Stojanovic 2017). By way of 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. Speaker A: The chili is tastier than the soup! **faultless**
Speaker B: No, the soup is tastier!
b. Speaker A: The Picasso is more beautiful than the Miró. **faultless**
Speaker B: No, the Miró is more beautiful.
c. Speaker A: Anna is taller than Zoe. **factual only**
Speaker B: No, Zoe is the taller of the two!

For the leading semantic approach to gradability, namely the degree-based analysis of Cresswell (1977), Kennedy (1997), Heim (2000) and others, ordering subjectivity is problematic. In such a framework, gradable adjectives lexicalize measure functions that map individuals to degrees on scales: *tall* is based on a HEIGHT measure function, *beautiful* on a BEAUTY function, and so forth (3). Comparative constructions are then analyzed as expressing relations between the degrees assigned to two individuals (4).

(3) a. $[\![\text{tall}]\!] = \lambda d \lambda x . \mu_{HEIGHT}(x) \succeq d$
b. $[\![\text{beautiful}]\!] = \lambda d \lambda x . \mu_{BEAUTY}(x) \succeq d$

(4) The Picasso is more beautiful than the Miró.
$\mu_{BEAUTY}(Picasso) \succ \mu_{BEAUTY}(Miro)$

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 a proposal 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,

2017; Umbach 2016; McNally and Stojanovic 2017). 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 dimensions of beauty might involve line, color, balance, and so forth. Subjectivity is proposed to arise because different individuals may weight these component dimensions differently, potentially resulting in a reversal of the relative ordering of two individuals. Exploring this line of explanation will prompt us to take a closer look at what it means for an adjective to be characterized as multidimensional.

Whichever approach one chooses to pursue, a crucial step in developing an adequate formal theory of ordering subjectivity (or subjectivity more generally) is to clarify which gradable adjectives are interpreted subjectively in 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 adjectives such as *tall* in that they lack commonly used measurement units. But they also different from those such as *beautiful* and *tasty* in that they appear to describe physical properties of objects in the world, rather than judgments based on internalized experiences. Can two individuals disagree faultlessly about which of two shirts is dirtier? which of two surfaces is rougher? which of two knives is sharper? As intuitions here are shaky, these questions were pursued experimentally, with the finding that ordering subjectivity is more widespread than has been previously recognized, and furthermore that in this respect, gradable adjectives pattern into not two but three subgroups: objective, subjective and mixed.

The primary proposal that is developed in this paper, which is based on the above two lines of investigation, is that there are two distinct sources of ordering subjectivity, namely multidimensionality and judge dependence. This proposal is formalized within a semantic framework in which gradable adjectives lexicalize not a single measure function but rather a set of such functions indexed to contexts and in some cases judges. Constraints on this set determine whether their comparative forms can be interpreted objectively, subjectively or in both ways. An ancillary conclusion that emerges is that adjectival multidimensionality is not a homogeneous phenomenon but rather has several distinct subtypes.

The structure of the paper is as follows: Sect. 2 presents the experiment and discusses some related phenomena. Section 3 briefly reviews existing semantic theories of subjectivity, with a view to assessing how well they are able to account for the experimental findings. Section 4 delves into the phenomenon of multidimensionality, offering evidence for its heterogenous nature. Section 5 presents the formal proposal, and Sect. 6 concludes.

2 Experiment: Faultless Disagreement Paradigm

The present study employs a novel faultless disagreement paradigm to diagnose the presence of ordering subjectivity among a wide range of gradable adjectives, with the goal of establishing a firmer empirical basis for formal semantic theories of the phenomenon.

2.1 Participants

Participants were 91 native speakers of English, recruited via the online participant marketplace Amazon Mechanical Turk (MTurk). Recruiting was limited to MTurk workers with U.S. IP addresses. Native language was confirmed via a question at the end of the survey; no participants were excluded on the basis of this question.

2.2 Materials

Stimuli were based on 35 gradable adjectives, which were divided into the following categories according to their status as dimensional versus evaluative, as well as the type of interpretation of the adjective in its positive form and the corresponding structure of the scale it lexicalizes[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 scales closed on both ends (**ABS2**): *full, empty*
- Absolute gradable adjectives with scales closed on one end (**ABS1**): *wet, dry, straight, curved, rough, smooth, clean, dirty, salty*
- 'Evaluative' adjectives (**EVAL**): *good, bad, beautiful, pretty, ugly, easy, interesting, boring, tasty, fun, intelligent, happy, sad*

Adjectives were assigned to these categories based on tests described in the literature, as follows. Relative gradable adjectives were identified as those for which both the

[1] In work on the semantics of gradable adjectives, it is now common to distinguish between context-dependent **relative** gradable adjectives and (more) context-independent **absolute** gradable adjectives (Kennedy and McNally 2005; Kennedy 2007). This distinction is proposed to derive from the structure of the scale lexicalized by the adjective: members of the absolute class have scales with maximum and/or minimum points, with these providing the standard for the adjective in its positive form, while members of the relative class have scales that are open on both ends, necessitating a contextual standard. A secondary objective of the present experiment was to explore the correlation between subjectivity and the relative/absolute distinction. Findings in this area are reported in Solt (2016), and due to space considerations will not be discussed here.

adjective and its antonym are acceptable in the frame *x is Adj but y is Adj-er*, and for which neither adjective nor antonym allows modification by *slightly*. Absolute gradable adjectives were identified as those for which either adjective or antonym is infelicitous in the above frame and/or can co-occur with *slightly*. Within the latter class, the division into doubly versus singly closed scales (ABS2 vs. ABS1) was based on judgments reported in the literature. An adjective was considered to have a numerical measure if its comparative form can be modified by a measure phrase. The evaluative category was selected to include adjectives of the sort discussed in the literature under the terms 'evaluative' (see especially Bierwisch 1989) or 'predicate of personal taste' (Lasersohn 2005 and many others). This is a mixed class, encompassing value, taste and aesthetic judgments, emotion words, and psychological predicates; its members are united, and distinguished from those of the other four categories, in that they do not denote external physical properties.

For each adjective, one or more dialogues were created, each featuring a disagreement between two speakers. For example:

(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: 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 vase on the table is more beautiful than the one on the bookshelf.
B: No, the vase on the bookshelf is more beautiful.

Adjectives were split across 4 lists, which were tested sequentially. Some adjectives occurred on more than one list, in different dialogue contexts. Each list contained 8-12 test items and 12 fillers. Fillers were split equally between two types: (i) those expressing factual disagreements (example: A: The judge found Frank guilty. B: No, the judge found Frank innocent.); (ii) those expressing differences of opinion, including statements based on vague nominal predicates (e.g. *jerk*), deontic and epistemic modals, statements of likelihood, and moral judgments. Sample size was 20–25 per list. See the Appendix for the full list of critical items.

2.3 Procedure

The study was executed online via Amazon MTurk, and 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. The task was introduced as follows:

(8) This study is about disagreements between people. Sometimes when two people disagree, only one of them can be right, and the other must be wrong. For example, in this short dialogue, Speaker A and Speaker B can't both be right, because Rosa can't have been born in both July and April.

> Speaker A: Rosa was born in July.
>
> Speaker B: No, Rosa was born in April.
>
> But sometimes when people disagree, there is no right or wrong answer - it's just a matter of opinion. Here's an example:
>
> Speaker A: Susan looks a lot like her sister.
>
> Speaker B: No—they don't look alike at all!
>
> In this HIT, you will see a series of short dialogues between two speakers A and B. Your task is to say whether there is a right or wrong answer, or whether it's a matter of opinion. Please answer based on your intuitions; do not think too long about each question.

Participants were then presented with a list of test and filler dialogues in pseudo-random order; their task was to classify each using one of two response options: "only one can be right; the other one must be wrong" and "it's a matter of opinion". The first of these was coded as a judgment of FACT; the second as a judgment of OPINION.

At the end of the questionnaire, participants were asked age and native language(s), and were given an opportunity to comment on the task. Participants were paid $0.75 for participation.

2.4 Results

The proportion of FACT judgments for each individual adjective and for the five subclasses in aggregate are displayed in Fig. 1. A mixed effect logistic regression 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 reference level was RELNUM.

Significant differences were found between RELNUM and ABS1 ($z = -7.016$, $p < 0.001$), RELNO ($z = -8.208$, $p < 0.001$) and EVAL ($z = -12.127$, $p < 0.001$). The difference between RELNUM and ABS2 was not significant ($z = -1.242$, $p = 0.214$). Among the classes found to differ significantly from RELNUM, subsequent post hoc testing via the *multcomp* package (Hothorn et al. 2008) using Tukey correction for multiple comparison found the following significant differences: ABS1 versus EVAL (z-ratio $= 11.049$, $p < 0.001$), RELNO versus EVAL (z-ratio $= 9.054$, $p < 0.001$) and ABS1 versus RELNO (z-ratio $= 3.803$, $p < 0.01$). Regarding the last contrast, however, an examination of the results for individual adjectives shows no clear separation between the two classes (see Fig. 1), suggesting that the overall difference found might be an artifact of the specific adjectives tested.

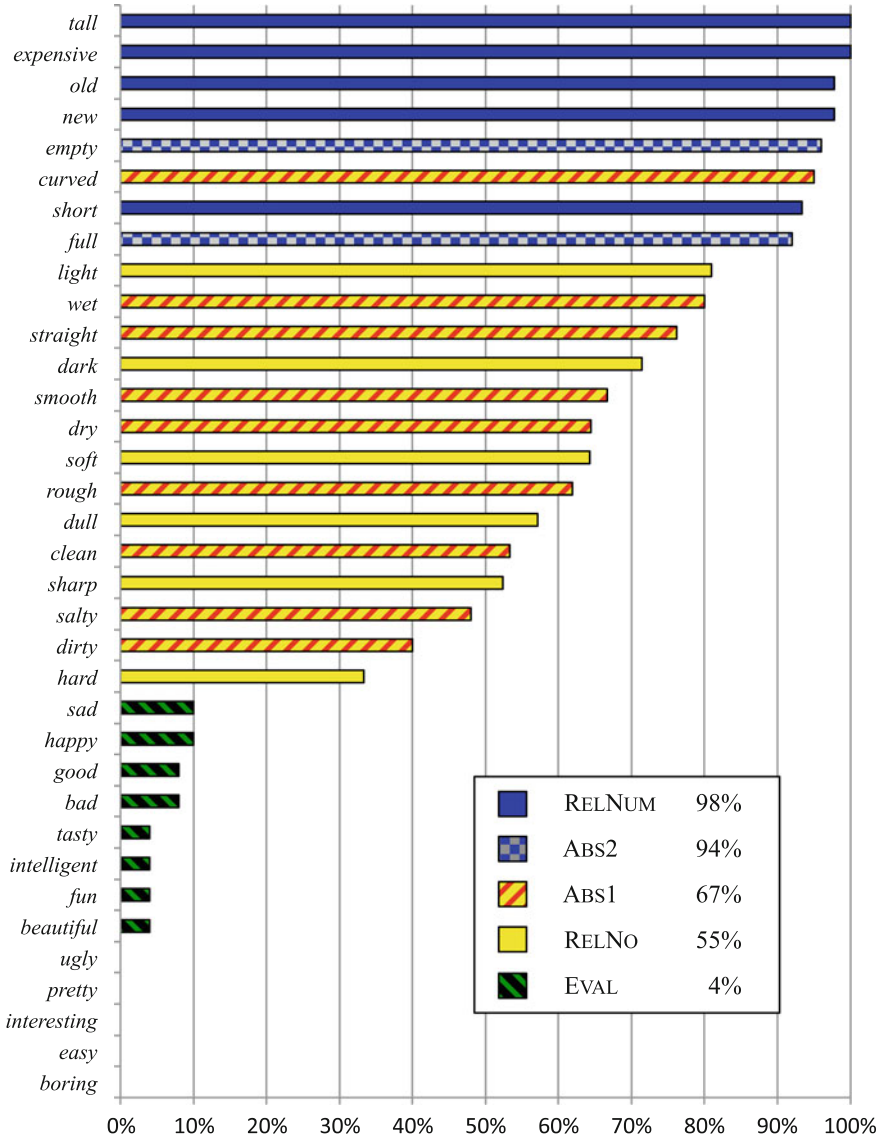


Fig. 1 Results of experiment—percent 'FACT' judgments

2.5 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 double-closed scale pair *full/empty* might be assimilated to this group, in that degrees of fullness (or emptiness) can be quantified in percentages (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 singly 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: objective, subjective and mixed. 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 dividing lines 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 objective group might in certain contexts allow subjective interpretations of their comparative forms, or conversely that members of the subjective class might in the right sort of contexts allow objective readings. However, one previously unrecognized finding appears quite clear: there is a large group of adjectives for which the interpretation of the comparative form is 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. The most obvious 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.

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 to members of the

intermediate group as well (see (9)–(11)). But when we turn to precise expressions of proportion such as *2.3 times as*, the picture changes (see (12)–(14)): these are possible for dimensional adjectives, and quite comically infelicitous for members of the evaluative class; for the mixed 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:

(9) a. The Eiffel Tower is twice as tall as the Great Pyramid.
b. The laptop is five times as expensive as the tablet.

(10) a. The Serta mattress is twice as hard as the Sealy mattress.
b. The blue shirt is five times as dirty as the green one.

(11) a. Anna is twice as beautiful as Zoe.
b. The roller coaster was ten times as fun as the ferris wheel.

(12) 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.

(13) a. ? The Serta mattress is 1.9 times as hard as the Sealy mattress.
b. ? The blue shirt is 5.1 times as dirty as the green one.

(14) 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.

The remainder of this paper is devoted to developing an account of these patterns. The next section briefly reviews existing semantic theories of subjectivity, focusing on their ability to explain the experimental results. One important proposal to come out of this work is that of multidimensionality as a source of subjectivity, particularly ordering subjectivity; this topic is explored in the section that follows.

3 Theories of Subjectivity

Adjectival subjectivity is the topic of a large body of research in formal semantics. The earliest of this work focused on predicates of personal taste such as *tasty* and *fun*, and pursued the general approach of accounting for their subjectivity by relativizing the interpretation of the adjective to a judge whose opinion or perspective is expressed. Theories in this area can be divided into two broad classes, which differ in 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 (15a). The contextualist approach (Stojanovic 2007; Sæbø 2009), by contrast, assumes that predicates of this sort feature an additional judge or experiencer argument (15b).

(15) a. $[\![\text{tasty}]\!]^{w,t,j} = \lambda x.x$ tastes good to j in w at t
b. $[\![\text{tasty}]\!]^{w,t} = \lambda y \lambda x.x$ tastes good to y in w at t

Elaborations on and combinations of these two approaches are found in Stephenson (2007) and Bylinina (2014, 2017), among others, while authors including Moltmann (2010) have proposed analyses that do not rely on the notion of a judge.

In the form presented, neither of the formulas in (15) accounts for ordering subjectivity. *Tasty* is a gradable adjective, having comparative and superlative forms (*tastier*, *tastiest*) and allowing composition with degree modifiers (*rather*/*very*/*extremely tasty*). But the above analyses localize subjectivity at the level of the unmodified positive form, thus providing no explanation for the possibility of subjective judgments regarding the ordering of two entities along a dimension such as tastiness. This might however be remedied fairly simply, by starting with a gradable entry of the form in (3) and relativizing the measure function to a judge.

A more fundamental 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.

From a different perspective, earlier authors including Kamp (1975) and Klein (1980) observed that certain gradable adjectives (e.g. *clever*) are dependent on multiple underlying dimensions for their ascription, one consequence of which is variability in judgments about the relative ordering of two individuals. More recent work (see especially Sæbø 2009; Kennedy 2013; Bylinina 2014; McNally and Stojanovic 2017; Umbach 2016) has connected this insight to the topic of subjectivity.

A central observation that has come out of this later work is that a wide range of gradable adjectives are subjective in their positive forms, including not only classical personal taste predicates but also other evaluative adjectives as well as vague gradable adjectives more generally; but only the first two of these are also subjective in their comparative forms (see again (1) vs. (2)). The conclusion is that there are two distinct loci of subjectivity. For vague gradable adjectives such as *tall*, subjectivity is localized not in the lexical meaning of the adjective itself but rather in the semantics of the positive morpheme *pos* that provides the threshold of applicability for the adjective in its unmodified form. For adjectives such as *tasty*, *fun* and *beautiful*, it derives from the lexical semantics of the adjective.

Kennedy (2013) 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." 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 there are also other possibilities).

The notion of multidimensionality as a source of subjectivity is taken up further by McNally and Stojanovic (2017) in the context of an investigation of aesthetic adjectives such as *beautiful*. They observe that "[d]eciding whether an adjective describing a multidimensional property holds of some individual involves not only determining a threshold of applicability but also determining the relative weight of each of the dimensions that contribute to the property in question. Here, again, there will be room for disagreement between speakers" (2017, p. 21). And further: "Two speakers may disagree about whether Ayumi is healthier than Mihajlo because they may disagree about whether one component of health or another (e.g. the state of the cardiovascular system vs. the immune system) should carry more weight" (2017, pp. 21–22). Multidimensionality is however only one source of subjectivity, others being experiential semantics (characterizing adjectives such as *tasty* and *interesting*) as well as evaluativity in the sense of expressing an attitude of positive or negative evaluation on the part of the speaker (e.g. *good*, *bad*, *beautiful*).

Bylinina (2014) proposes a formal analysis of adjectival subjectivity that explicitly incorporates multidimensionality. Her account is based in part on the observation that the class of adjectives exhibiting ordering subjectivity can itself be further subdivided: subjective readings for the comparative are possible for both adjectives such as *fun*, *tasty* 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:

(16) 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. Her formalization is the following (where Q is a dimension contributing to intelligence, w_Q^j is the weight assigned by j to Q, $m_{x,Q}$ is the measure of an individual x with respect to Q and s_Q is the standard of applicability for Q).

(17) $[\![m_{x,\textbf{intelligent}}]\!]^{c;w,t,j} = \lambda x.\sqrt{\sum_{Q}[w_Q^j(m_{x,Q} \succ s_Q)]^2)}$

Umbach (2016) takes a somewhat similar approach, analyzing the evaluative adjective *beautiful* in terms of a generalized measure function that maps entities to points in a multidimensional attribute space.

In summary, several authors have argued convincingly that a source of adjectival subjectivity, and specifically ordering subjectivity, is the multidimensional nature of the properties in question. But note that each of these accounts has treated multidimensionality-based subjectivity as a variety of judge dependence: two judges may weight an adjective's dimensions differently, potentially giving rise to disagreements about orderings. This brings up a more general point. In all of the works discussed in this section, the focus has been on 'subjectivity' in the sense of the diverging perspectives of distinct speakers. This perhaps stems from the initial focus on personal taste predicates such as *tasty* and *fun*, which so clearly express individuals' judgments or tastes. When we expand our focus to the full range of adjectives considered in the present work, it becomes clear that differences between judges are not the only source of variable judgments regarding orderings; rather, it seems that a single speaker's judgments may also be potentially uncertain or changeable. Consider for example two shirts, one which is clean except for a grass stain on the sleeve, the other slightly dingy overall. Which one should 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. Variability of this sort cannot be accounted for by relativization to a judge, but rather seems to reflect a more general sort of context dependence.

In the next section, I take a more in-depth look at the nature of adjectival multidimensionality. This will form the basis for the formal account in Sect. 5, which also seeks to clarify the relationship between multidimensionality and judge dependence.

4 Identifying Multidimensionality

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. This turns out to be less straightforward than it might initially seem.

4.1 Sassoon's Theory of Multidimensionality

As noted above, it has long been recognized that some gradable adjectives are multidimensional (see especially Kamp 1975 and Klein 1980; for discussion of multidimensionality more broadly, see also Bartsch and Vennemann 1972; Bartsch 1984, 1986; Landman 1989). But 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 as *healthy*, the default binding operator is universal quantification: to be healthy is to be healthy in all contextually relevant respects (18a). For disjunctive adjectives such as *sick*, the default is existential quantification: to be sick is to be sick in some relevant respect(s) (18b). Adjectives such as *intelligent* are mixed, with pragmatics determining the binding operation.

(18) a. *healthy*: $\lambda x.\forall Q \in DIM(healthy) : Q(x)$
b. *sick*: $\lambda x.\exists Q \in DIM(sick) : Q(x)$

Comparatives might 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 particular 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* (19) or inquired about via a *wh-phrase* (20); dimensions may be quantified over (21); and quantificational force may be restricted by exception phrases (22).[2] None of these are possible with (uni-)dimensional adjectives such as *tall*.

(19) 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.

(20) a. In what respects is the patient healthy/sick?
b. In what respects are the boxes identical?
c. #? In what respect is Zoe tall?

(21) 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).

(22) a. The patient is healthy/not sick except for high blood pressure/asthma/a slight cold.

[2]Which 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 these details here.

b. The boxes are identical except for size/color.
c. # Zoe is tall except for …

Sassoon backs up these judgments with extensive corpus and experimental data, particularly relating to the pattern in (22).

Multidimensionality of the sort described here has also been proposed to play a role in other linguistic patterns, such as the acceptability of so-called borderline contradictions (see Égré and Zehr, this volume).

4.2 Varieties of Multidimensionality

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*. More generally, when we look at the mixed and purely subjective groups that emerged from the experiment, we see that many 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 *salty* can be put in this class: while one might think that degree of saltiness is dependent on a single dimension, 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 the multidimensional status of such adjectives via tests based on the constructions in (19)–(22), and thereby clarify which of the adjectives exhibiting ordering subjectivity are multidimensional, the results are quite mixed. Consider to start the personal taste predicates *tasty* and *fun*, both of which patterned as purely subjective in our experiment:

(23) 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/??

(24) 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/??

Compared to the corresponding examples with *healthy*, *sick* and *identical*, it seems more difficult to continue the sentences in (23a), (24a), or to answer the questions

in (23b), (24b).[3] What are the respects of tastiness and fun that contribute to the attribution of these predicates? If anything, the questions seem to favor a rhetorical interpretation, challenging the interlocutor to name even one ground for calling the chili tasty or the roller coaster fun. Similarly, universal and existential quantification over dimensions is moderately acceptable ((23c), (24c)), producing emphatic and hedging effects, respectively, but precise counting of dimensions (??*fun/tasty in three respects*) is rather odd. Finally, it is certainly 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*); but after these the task becomes more difficult (see (23d), (24d)), suggesting that there is a considerable residual meaning that cannot be easily separated into discrete dimensions.

A similar issue emerges with other evaluative predicates, where we see that even when examples parallel to (19)–(22) sound felicitous, they do not necessarily involve specification of or quantification over dimensions. Take for example *beautiful*, another 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 (25), where the listed aspects seem to be not component dimensions of the predicate *beautiful* but rather component parts of a complex entity or event that is the subject of predication.

(25) a. The wedding was beautiful in every respect … the weather, the venue, the bride's dress, and most of all, the people!

b. This newly built home is beautiful in every way, featuring a welcoming great room with stone fireplace, a light-filled open-plan kitchen, and a spacious master bedroom suite.

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 2017). 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 ...*). 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 really exploring the difficulty of creating true dimensional examples for those such as *beautiful*.

[3]For myself, examples of this sort are quite bad; a reviewer, however, found them more acceptable. Such between-speaker variation is itself indicative of the difficulty in classifying an adjective as multidimensional versus unidimensional.

Here I do not mean to claim that adjectives such as *tasty*, *fun* and *beautiful* can never have a multidimensional interpretation (in Sassoon's sense); the possibility of dimensional exception phrases and the like is enough to show this cannot be right. The multidimensional interpretation might in particular be more available to experts in the relevant domains (think for example of a food writer or art critic), who have a trained ability to introspect into the factors underlying their judgments. The point is rather that such adjectives, while without doubt multidimensional at the conceptual level, also have an interpretation—perhaps the most salient one—on which they behave grammatically as if they were unidimensional.[4]

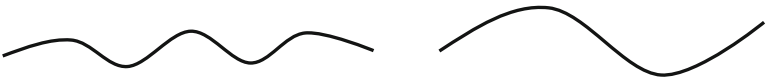
Consider now the adjectives in our mixed group. Of these, *clean* and *dirty* are discussed as multidimensional by Sassoon, and this is supported by the above-described tests:

(26) a. In what respect(s) was the shirt clean/dirty?
b. The shirt was clean/dirty in 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, the results are quite different. Taking *except* phrases as an example, it is difficult to construct true dimensional completions of examples such as the following:

(27) 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 …

Yet 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.

(28) 

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 second with an all-over 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

[4]I thank the reviewers for pointing out the need to clarify this point.

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 is not 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 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 are not entirely felicitous in the constructions in (19)–(22). The reverse is also true: certain adjectives that are generally considered to be dimensionally ambiguous rather than multidimensional are relatively acceptable with *respect*. Examples are *large* and *long*:

(29) In which respect is London larger than New York?
Land area ✔ Population size ✘

(30) The sofa is larger than the bench in every respect.

(31) a. The trip to Tübingen is longer than the trip to Konstanz.
b. In which respect—travel time or distance in kilometers?

This suggests that *which respect* questions at least might in fact offer a test for the contextual dependence of the communicated dimension, rather than for multidimensionality.

In summary, the preceding discussion suggests that adjectival multidimensionality is not a homogenous phenomenon. There are gradable adjectives such as *healthy* and *identical* that are multidimensional in what might be called a quantificational sense: their component dimensions are readily named, easily separated, and grammatically active, and for the positive form of the adjective at least, a variety of tests suggest that they are integrated by means of quantificational operators. But there are other sorts of intuitively multidimensional adjectives—examples being *bumpy*, *curved*, *salty* and (in my judgments) *fun* and *tasty*—for which the individual component dimensions are much less grammatically, or even conceptually, accessible. The attribution of such predicates certainly depends on multiple aspects or properties of the object of predication; but (ordinary) speakers are quite likely not aware of or able to name these aspects and properties. Furthermore, that such adjectives tend to pattern as unidimensional rather than multidimensional on the above-described tests suggests that their dimensions do not compose via universal or existential quantification but rather are integrated in some other manner to create a single, complex dimension. The dividing line between these two variants of multidimensionality is not entirely sharp; quite plausibly, some adjectives (e.g. perhaps *beautiful*) allow both sorts of

[5]The 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.

interpretations, or combine the two on a single usage. Given this, I will continue to use the term 'multidimensional' to describe both sorts of adjectives.

For the purposes of the present paper, the crucial observation is that both varieties of multidimensionality—the quantificational variety and the complex dimension variety—appear to give rise to the possibility of subjective judgments regarding orderings. Capturing this observation is a central goal of the formal analysis proposed below.

4.3 Multidimensionality and Evaluation

There is a further distinction among the class of adjectives that are multidimensional in the broad sense, which is subtle but I believe nonetheless real, and which is relevant to the adequate formal analysis of such adjectives.

For classic examples of multidimensional adjectives such as *healthy/sick* and *identical* as well as those such as *clean/dirty*, *straight/curved* and *flat/bumpy*, the overall meaning of the adjective is in a sense built up directly from its component dimensions, integrated in some contextually determined way. The degree of sickness of an individual is determined by the nature and perhaps severity of his relevant illnesses; the bumpiness of a road by the size/shape/etc. of the bumps and dips on it; the straightness or curvedness of a line, by the number or shape or other mathematical properties of the curves on it.

For so-called evaluative adjectives, namely those of the sort that made up the EVAL group in the present experiment, there is something more that this. Specifically, while the adjective's meaning is based in some way on multiple underlying properties of the object of predication, there is also an inherent human element. Some are experiential in nature, as diagnosed by the possibility of modification by an experiencer PP (e.g. *tasty to me*; *fun for me*; see Sect. 3); experiential meaning requires an experiencer. Others express an aesthetic or taste judgment. Yet others convey an emotion, and are thus necessarily rooted in the perceptions or feelings of an individual. And while it is arguably not an inherent aspect of their meaning, on their typical uses most are evaluative in the sense of expressing a positive or negative value judgment; value judgments (like taste and aesthetic judgments) require an individual who judges. To borrow a term used by McNally and Stojanovic (2017), all of these sorts of adjectives require the "intermediation of a sentient individual" in their attribution.

The claim that I would thus like to make is that multidimensional adjectives can stand in two distinct types of relations to their component dimensions. For those such as *healthy*, *clean/dirty* and *flat/bumpy*, the adjective's overall meaning can be expressed directly as a function of its dimensions (though the function is context dependent, and might not be fully transparent to the ordinary speaker). 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 above claim is similar to one made by the moral philosopher Hare (1952), who 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'. 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 (see also Umbach 2016 for related discussion).

It is rather difficult to design diagnostics for the distinction suggested above, but a possible one is based on follow-up questions. For at least some adjectives of the *healthy/clean/bumpy* sort, a speaker can be asked to clarify her assertion by means of a *what respect/way* question.

(32) a. Fred is healthier/sicker than Tom.
b. The blue shirt is cleaner/dirtier than the green one.
c. Weserstrasse is bumpier that 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, a more natural way to question the speaker's assertion is to ask for her reasons for it, for example with *What makes you say that?*

(33) 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?

This suggests a recognition that for adjectives of the latter sort, the objective properties of the subject(s) of predication contribute to the attribution of the adjective only indirectly, through their effect on the perceptions or judgments of the speaker.

4.4 Summary

We have seen here that a wide variety of gradable adjectives are multidimensional in a conceptual sense, being dependent on multiple properties of an object for their

attribution, and thereby distinguishable from straightforward (uni-)dimensional adjectives, which lexicalize a single, typically measurable dimension. But the multidimensional class can itself be further subdivided. In some such adjectives (or perhaps more accurately, uses of such adjectives), the component dimensions are readily accessible and grammatically active, while in others they are integrated in a way that is not transparent to the average speaker. And I have argued that the meaning of some conceptually multidimensional adjectives can be expressed as a direct function of their dimensions, while for others, their dimensions play a more indirect role in their meaning, as factors contributing to some sort of judgment by a sentient individual. Importantly, all of these varieties of multidimensionality result in ordering subjectivity, though I will propose that they do so in different ways.

5 Proposal

In this section, I outline a theory of gradable adjective meaning that formalizes the observations from the prior two sections, and that provides the basis for explaining the availability of objective and subjective 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:

(34) $S = \langle D, \succ, DIM \rangle$, where

- DIM is a dimension of measurement
- D is a set of degrees
- $\succ$ is an ordering relation on D

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 a domain of measurement Dom (e.g. the domain of individuals or of events, or a subset thereof) to some scale S tracking dimension DIM.

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

(35) $[\![\text{Adj}]\!]^c = \lambda d \lambda x . \mu^c_{DIM}(x) \succeq d$

To put this differently, gradable adjectives on this view lexicalize dimensions rather than particular scales or measure functions. A dimension is a property that an entity can have more or less of. A measure function corresponding to that dimension is a mapping from individuals to degrees that represent the extent that each individual has the property in question. As a very simple example to demonstrate that these two things are not equivalent, the single dimension *HEIGHT* may be tracked by a function that maps individuals to their height in inches, or alternately by a function that maps individuals to their height in centimeters. For a simple unidimensional adjective such as *tall*, this might be the only sort of variation that is possible; but for other classes, there are further possibilities. Below I will argue that the availability of objective versus subjective readings for the comparative derives from constraints on the possible variation in the family of functions $\{\mu^c_{DIM} : c \in C\}$ that is the semantic content of 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 the set of measure functions lexicalized by the adjective 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.

There are several routes to such a mapping. 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 (see Krifka 1989; Sassoon 2010; Lassiter 2011 and references therein). 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. Other dimensions that satisfy additivity include weight, depth, width, length, volume and duration. Even 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. Additivity provides the possibility of numerical measurement: some standard element is selected as the basis of a unit of measurement, and the measure of any individual can then stated in terms of multiples of this standard. A 6-meter-tall tree, for example, is one whose height is equivalent to the concatenation of six copies of a 1-meter standard element.

Formally, additivity may be encoded via a constraint on the set of measure functions $\{\mu^c_{DIM} : c \in C\}$ that is lexicalized by the adjective. For readers interested in the

technical details, the constraint is that in (36) (where $\oplus$ is the relevant concatenation operation). A sample denotation for an adjective satisfying this constraint is (37).

(36) **Additive measure functions:**
$\forall c \in C$ and $\forall a, b \in Dom,\ \mu^c_{DIM}(a \oplus b) = \mu^c_{DIM}(a) + \mu^c_{DIM}(b)$

(37) $[\![\text{tall}]\!]^c = \lambda d \lambda x . \mu^c_{HEIGHT}(x) \succeq d$,
where $\forall c \in C$ and $\forall a, b \in Dom$,
$\mu^c_{HEIGHT}(a \oplus b) = \mu^c_{HEIGHT}(a) + \mu^c_{HEIGHT}(b)$

Beyond additivity, there are other possible routes to numerical measurement. First, there are dimensions for which **natural, speaker-external phenomena** serve as the basis for measurement units. 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 that are **derivable from measurable dimensions 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. Other dimensions in this class might be purity (defined as volume of impurities relative to total volume) and speed (distance traveled divided by duration). In each of these cases, numerical measures can be derived on the basis of the component measure functions, which enables proportional or ratio measure expressions, as in *20% full*, *90% pure*, and *5 kilometers per hour faster/slower*.

Formally, adjectives falling in this class are those that satisfy the constraint in (38). As an example, the corresponding lexical entry for the adjective *full* is given in (39):

(38) **Context independent derived measure functions**:
$\forall c \in C$ and $\forall x \in Dom$,
$\mu^c_{DIM}(x) = f(\mu^c_{DIM_1}(x), \mu^c_{DIM_2}(x), \ldots, \mu^c_{DIM_n}(x))$,
where $\mu^c_{DIM_1},\ \mu^c_{DIM_2}, \ldots \mu^c_{DIM_n}$ are objective measure functions

(39) $[\![\text{full}]\!]^c = \lambda d \lambda x . \mu^c_{FULLNESS}(x) \succeq d$,
where $\forall c \in C$ and $\forall x \in Dom$,
$$\mu^c_{FULLNESS}(x) = \frac{\mu^c_{VOLUME}(content(x))}{\mu^c_{VOLUME}(capacity(x))}$$

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* and *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 adjectives whose comparative forms can be interpreted subjectively, as diagnosed by the possibility of faultless disagreement. The overall approach that I pursue is that **ordering subjectivity arises when the set of measure functions lexicalized by the adjective is such that a difference in context can result in a difference in the relative ordering of two individuals**. Building on the previously discussed observations by Bylinina (2014), as well as the discussions in Sects. 3 and 4, I propose that this can come about in two ways, namely through multidimensionality and dependence on a judge.

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 Bylinina (2014) in proposing that adjectives of this sort are associated in each context c with a set of component dimensions $DIM_1^c, DIM_2^c, \ldots, DIM_n^c$. Departing somewhat from these authors, I further assume that to each dimension DIM_i^c there corresponds a measure function $\mu^c_{DIM_i^c}$, 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. (40) specifies the form of such functions, and (41) gives a plausible if undoubtedly overly simplistic entry for *dirty* in this form.

(40) **Context-dependent derived measure functions**:
$\forall c \in C$ and $\forall x \in Dom$,
$\mu^c_{DIM}(x) = f^c(\mu^c_{DIM_1^c}(x), \mu^c_{DIM_2^c}(x), \ldots, \mu^c_{DIM_n^c}(x))$

(41) $[\![\text{dirty}]\!]^c = \lambda d \lambda x . \mu^c_{DIRTINESS}(x) \succeq d$,
where $\forall c \in C$ and $\forall x \in Dom$,
$$\mu^c_{DIRTINESS}(x) = \frac{\sum_{i=1}^{n} k_i^c \cdot \mu^c_{AMOUNT}(dirt_i^c(x))}{\mu^c_{SIZE}(x)}$$

Here the individual dimensions that underlie the adjective's meaning may themselves be objectively measurable. Subjectivity derives from the potential for variation in the choice of these dimensions and how they are combined.

Note that in the above formulation I have not made a distinction between the quantificational and complex dimension varieties of multidimensionality discussed in the previous section, though I leave open the possibility that this may ultimately prove necessary.

Judge dependence. The entries in (40) and (41) do not explicitly reference a judge. Rather, measure functions are indexed to contexts; 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 the earlier 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 necessarily involves a human element, or what was earlier called the mediation of a sentient individual. These include value judgments (*good/bad*), aesthetic judgments (*beautiful/ugly*), taste ascriptions (*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 build on 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 formally as follows:

(42) **Judge dependent measure functions**:
$[\![\text{Adj}]\!]^{c;j} = \lambda d \lambda x . \mu^{c;j}_{DIM}(x) \succeq d$
where $\mu^{c;j}_{DIM}(x)$ should be interpreted as
'the degree to which j judges x in context c to have property DIM'

(43) $[\![\text{beautiful}]\!]^{c;j} = \lambda d \lambda x . \mu^{c;j}_{BEAUTY}(x) \succeq d$

Again it is possible that this class must be further subdivided, for example to distinguish between adjectives with experiential semantics such as *tasty* and *interesting* and evaluative predicates such as *beautiful* (per McNally and Stojanovic 2017, cf. Sect. 3). I do not attempt to address this here.

The formulations in (42) and (43) do not represent adjectives such as *beautiful* as explicitly multidimensional. The rationale for this derives from the observations in Sect. 4. As was noted there, the dimensions underlying adjectives such as *beautiful*

and *tasty* are not as accessible grammatically or even conceptually as for paradigm cases such as *healthy*. More fundamentally, I argued in that section that dimensions play a different role for such adjectives than for those such as *healthy* and *clean/dirty*, being not direct 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 tentatively 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.

As further evidence for the need to distinguish between underspecification of meaning due to multidimensionality (as in (40)) and judge dependence (as in (42)), there are adjectives that appear to be ambiguous between the two types of subjective interpretations. Consider again the adjective *bumpy*, and the two city streets from our earlier example, one smooth 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 (44) is true, the issue being how to weight the different sorts of bumps and dips to arrive at an overall measure of bumpiness. But the disagreement in (45) implies something further, namely that the speakers have experienced the two streets in question (e.g. by riding a bike over them):

(44) a. Weserstrasse is bumpier than Friedelstrasse.
b. Friedelstrasse is bumpier than Weserstrasse.

(45) a. I find Weserstrasse bumpier than Friedelstrasse.
b. I find Friedelstrasse the bumpier of the two.

Thus *bumpy* appears to allow both a simple multidimensional interpretation and an experiential (and thus judge dependent) interpretation. The adjective *sharp* provides a similar example: on examining two knives under a microscope, one might be uncertain as to which is sharper, the issue again being how precisely degrees of sharpness should be determined. 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 turn now to one of the central findings from the present experimental research, namely that many of the adjectives tested—among them *hard/soft*, *sharp/dull*, *clean/dirty*, *rough/smooth* and *salty*—allowed both types of interpretations. The framework for adjectival semantics proposed in

this section is able to account for the existence of this mixed group, and also for the difference between these adjectives and members of the purely subjective group.

Mixed behavior may first of all arise as the result of an ambiguity between measurable/objective and subjective/experiential interpretations, a possibility suggested by Kennedy (2013). This does not require us to posit a lexical ambiguity for the adjectives in question; rather, it is already allowed for by the underspecified, context-dependent template for gradable adjective meaning in (35). 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 (and thus could potentially be associated with a numerical measure), 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*.

This is however not the only possible source of mixed behavior. Rather, the logical forms of adjectives of the multidimensional sort themselves allow objective as well as subjective interpretations. The intuition behind the formalisms in (40) and (41) is that the lexical entries of adjectives such as *clean* and *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, or when a single speaker is uncertain as to how to specify them. 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 $c \in C$, the order of the degrees assigned to these two shirts relative to the dimension dirtiness remains the same. As a special case of objective judgments with this class, the present account correctly predicts that a shirt that is completely free of dirt will necessarily be assessed as cleaner/less dirty than one with some amount of contextually relevant dirt: regardless of how dirt types are weighted in a particular context, the mathematical form of the lexical entry in (41) has the consequence that the former shirt will be mapped to the 0 point on the dirtiness scale, while the latter will be assigned some positive value.[6]

Returning to the experimental results, we might hypothesize that in the case of multidimensional adjectives such as *clean/dirty*, *smooth/rough*, *dull/sharp* and perhaps others, subjects who gave FACT and OPINION judgments made different assumptions about the two entities under consideration, or about the relevant context. The first group may potentially have assumed that the entities were different to

[6]I thank an anonymous reviewer for pointing out this example.

such a degree, or the context specified to such a degree, that all available measure functions would yield the same ordering. The second group may conversely have assumed that the individuals were sufficiently close in their properties, or that the context was sufficiently underspecified, that that there were available interpretations (i.e. measure functions) that would yield different orderings. I believe this is a plausible explanation for the experimental findings, though it would benefit from further experimental investigation.

Crucially, though, for judge-dependent adjectives of the *beautiful* and *tasty* 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, ..., DIM_n^c and the function f^c; for certain pairs of entities or situations of utterance, these constraints have the effect of eliminating the possibility of variation in orderings. 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. Thus drawing a distinction between multidimensionality and judge dependence as sources of ordering subjectivity further helps to explain why some adjectives that are subjective in this sense also allow objective interpretations, while others do not.

Table 1 Classes of gradable adjectives

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

5.5 Summary

The observations from this section are summarized in Table 1, which presents 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 form. 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 diagnostics to assign adjectives to the appropriate category or 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 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, i.e. a sentient individual whose judgments, tastes or emotions are expressed. Formally, I have developed this insight in a theory of gradability on which the availability of objective versus subjective readings of the comparatives derives from the formal properties of the measure functions lexicalized by gradable adjectives.

There are a number of important issues that I have not been able to address completely in the context of the present work. In particular, while the experiment reported here demonstrated the existence of two distinct classes of subjective adjectives, it does not provide direct evidence for the underlying distinction I have proposed, namely multidimensionality versus judge dependence. It is to be hoped that the predictions of this proposal can be tested more directly in future experimental research. Here, an issue is that it is challenging to find adequate diagnostics for multidimensionality

and especially judge dependence that might serve as the basis for an experimental paradigm (cf. the discussion in Sect. 4). With regards to diagnosing the presence of an explicitly represented sentient judge, the varying acceptability of different types of follow-up questions discussed in Sect. 4.3 might provide a starting point; another potential direction involves the possibility of 'coordination by stipulation' discussed by Kennedy and Willer (2016), which plausibly is sensitive to the distinction proposed in the present paper. Alternately, a promising more indirect approach is found in the work of Kaiser and Herron Lee (2017a, b), who show that predicates of personal taste and multidimensional adjectives pattern differently in the complement position of Experiencer-Theme verbs such as *hear*, which make salient an experiencer who may be associated with the corresponding role introduced by the adjective. This area is ripe for further research.

Looking more broadly, in focusing on the comparative I have made no attempt to address the interpretation of gradable adjectives of various sorts in their positive forms. One question that merits further investigation relates to the semantically multidimensional class. As was discussed above, there is considerable evidence that the positive forms of a subclass of such adjectives involve quantification over dimensions, something that is not easily expressed in the present formalization, in which explicitly multidimensional adjectives are analyzed in terms of derived measure functions. This suggests there remains work to be done in integrating the present findings with those from previous work on multidimensionality. I have furthermore not attempted to apply the present analysis to the subjectivity of the positive form of 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.

Acknowledgements I would like to express my gratitude to the editors of this volume Elena Castroviejo-Miró, Louise McNally and Galit Sassoon, as well as three anonymous reviewers, whose insightful comments and suggestions on earlier versions have led to a much improved paper. Thanks also to Carla Umbach, Nat Hansen, Kjell Johan Sæbø and the audiences at ExpGrad, SuB20 and the ZAS for very helpful discussion, and to Nadja Reinhold and Lisa Reimann for assistance with the experimental research. All errors are of course my own. This work was supported by the German Science Foundation (DFG) under grant SO1157/1-1.

Appendix

This appendix provides the full stimuli (critical items) used in the experiment.

List 1 A: Anna's apartment is dirtier than Paul's.
B: No, Paul's place is dirtier.

A: Frank is shorter than Jimmy.
B: No, Jimmy is the shorter one.

A: John and Fred look similar but John is a little taller than Fred.
B: No! Fred is the taller one of the two.

A: Lea and Marie are both sad but Marie is sadder.
B: No, Lea is sadder.

A: Lilian's car is newer than Noemi's car.
B: No, Noemi's is definitely the more recent one.

A: My painting is prettier than yours.
B: No! My painting is definitely prettier than yours.

A: Susan just got out of the water, but her hair is already drier than mine.
B: No, it's not - your hair is definitely drier than Susan's.

A: The green towel is wetter than the red one.
B: No, the red towel is wetter.

A: This building is older than the building Julia lives in.
B: No, Julia's building is older.

A: The mug is cleaner than the spoon.
B: The spoon is cleaner than the mug.

A: This cat is happier than that dog.
B: No, the dog is the happier one of them.

A: Those sneakers are uglier than the Converse sneakers you tried on earlier.
B: No, the Converse sneakers were uglier.

List 2 A: Can I borrow your pencil? Mine is duller than the one you have.
B: No, my pencil is even duller than yours.

A: Caryl and Tina both have blond hair, but Caryl's is lighter than Tina's.
B: No, Tina's hair is definitely lighter than Caryl's.

A: Give those kids the green ball to play with, it's softer than the red one.
B: No, the red ball is softer.

A: I would rather use the yellow pillow—it's harder than the white one.
B: No, the white pillow is the harder one of the two.

A: Math is easier than Geography.
B: Geography is a lot easier than Math!

A: Take the red knife, it's sharper than the one you're using.
B: No, the knife I have now is sharper than the red one.

A: The fence in front of Mr. Harington's house is straighter than the one in front of Mr. Rave's house.
B: No, Mr. Rave's fence is straighter.

A: The old Ipod Touch 4G is more expensive than the new Ipod Touch 5G.
B: No, the new one is the more expensive device.

A: The second line on that graph is more curved than the first one.
B: No, the first line is more curved than the second.

A: The walls in the dining room are darker than the walls in the living room.
B: No, the walls in the living room are darker.

A: This small piece of paper is smoother than that big piece of paper.
B: No, the big piece is smoother.

A: This stone right in front of us is rougher than that one in the back.
B: No, the stone in the back is rougher.

List 3 A: Frank is shorter than his friend Jimmy.
B: No, Jimmy is the shorter one.

A: John and Fred look similar but John is taller than Fred.
B: No, Fred is the taller one of the two.

A: Lillian and Nicole both have the same kind of cellphone, but Lillian's is newer than Nicole's.
B: No, Nicole's phone is newer than Lillian's.

A: Look—Tommy's shirt is dirtier than the one his little brother Billy is wearing.
B: No, Blly's shirt is dirtier than Tommy's.

A: My apartment building is older than the building Julia lives in.
B: No, Julia's building is older.

A: Susan just got out of the water, but her bathing suit is already drier than mine.
B: No, it's not—your bathing suit is drier than Susan's.

A: The green towel on the hook is wetter than the blue one.
B: No, the blue towel is wetter.

A: The lecture we heard last week was more boring than today's lecture.
B: No, today's lecture was more boring.

A: The mug you just handed me is cleaner than the one on the counter.
B: No, the one on the counter is cleaner.

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: The program we watched about India was more interesting than the one about Japan.
B: No, the program about Japan was the more interesting of the two.

A: The vase on the table is more beautiful than the one on the bookshelf.
B: No, the vase on the bookshelf is more beautiful.

List 4 A: I just read the essay John wrote and it is worse than Bill's.
B: No it isn't. The one Bill wrote is worse.

A: Look at Sue's new bike—it's better than Anne's.
B: No, Anne's bike is better.

A: The cream cake is tastier than the chocolate cookies.
B: No, the chocolate cookies are tastier.

A: The math professor is more intelligent than the physics professor.
B: No, I disagree. The physics professor is the more intelligent one.

A: The movie theater is emptier today than it was yesterday.
B: No it wasn't. It was emptier yesterday.

A: The roller coaster was more fun than the ferris wheel.
B: No it wasn't! The ferris wheel was more fun.

A: The vegetable soup is saltier than the chicken soup.
B: No, the chicken soup is saltier.

A: The wine bottle is fuller than the champagne bottle.
B: No, the champagne bottle is fuller.

References

Bartsch, R. (1984). The structure of word meanings: Polysemy, metaphor, metonymy. In F. Landman & F. Veltman (Eds.), *Varieties of formal semantics* (pp. 25–54). Dordrecht: Foris.

Bartsch, R. (1986). Context-dependent interpretations of lexical items. In J. Groenendijk, D. de Jongh, & M. Stokhof (Eds.), *Foundations of pragmatics and lexical semantics*, GRASS 7 (pp. 1–26). Dordrecht: Foris.

Bartsch, R., & Vennemann, T. (1972). The grammar of relative adjectives and of comparison. *Linguistische Berichte*, *20*, 19–32.

Bates, D. M., Mächler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using 'Eigen' and S4. R package version 1.1-7. http://CRAN.R-project.org/package=lme4.

Bierwisch, M. (1989). The semantics of gradation. In M. Bierwisch & E. Lang (Eds.), *Dimensional adjectives: Grammatical structure and conceptual interpretation* (pp. 71–261). Berlin: Springer.

Bylinina, L. (2014). *The grammar of standards: Judge-dependence, purpose-relativity, and comparison classes in degree constructions*. Dissertation, Utrecht University.

Bylinina, L. (2017). Judge-dependence in degree constructions. *Journal of Semantics*, *34*(2), 291–331.

Christensen, C. M. (1980). Effects of solution viscosity on perceived saltiness and sweetness. *Perception and Psychophysics*, *28*(4), 347–353.

Cresswell, M. J. (1977). The semantics of degree. In B. H. Partee (Ed.), *Montague Grammar* (pp. 261–292). New York: Academic Press.

Hare, R. M. (1952). *The language of morals*. Oxford: Clarendon Press.

Heim, I. (2000). Degree operators and scope. In J. Brendan & T. Matthews (Eds.), *Proceedings of the 10th Semantics and Linguistic Theory Conference (SALT 10)* (pp. 40–64). Ithaca, NY: CLC Publications.

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, *50*(3), 346–363.

Kaiser, E., & Lee, J. H. (2017a). Predicates of personal taste and multidimensional adjectives: An experimental investigation. In *Proceedings of the 35th West Coast Conference on Formal Linguistics (WCCFL35)*. Somerville, MA: Cascadilla Proceedings Project.

Kaiser, E., & Lee, J. H. (2017b). Experience matters: A psycholinguistic investigation of predicates of personal taste. In D. Burgdorf, J. Collard, S. Maspong, & B. Stefánsdóttir (Eds.), *Proceedings of Semantics and Linguistic Theory 27 (SALT 27)*. Washington, DC: Linguistic Society of America.

Kamp, H. (1975). Two theories about adjectives. In E. L. Keenan (Ed.), *Formal semantics of natural language* (pp. 121–155). Cambridge: Cambridge University Press.

Kennedy, C. (1997). *Projecting the adjective: The syntax and semantics of gradability and comparison*. Dissertation, University of California at Santa Cruz.

Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy*, *30*(1), 1–45.

Kennedy, C. (2013). Two sources of subjectivity: Qualitative assessment and dimensional uncertainty. *Inquiry*, *56*(2–3), 258–277.

Kennedy, C., & McNally, L. (2005). Scale structure, degree modification and the semantics of gradable predicates. *Language*, *81*(2), 345–381.

Kennedy, C., & Willer, M. (2016). Subjective attitudes and counterstance contingency. In M. Moroney, C.-R. Little, J. Collard, & D. Burgdorf (Eds.), *Proceedings of Semantics and Linguistic Theory 26 (SALT26)* (pp. 913–933). Washington, DC: Linguistic Society of America.

Klein, E. (1980). A semantics for positive and comparative adjectives. *Linguistics and Philosophy*, *4*(1), 1–45.

Kölbel, M. (2004). Faultless disagreement. In *Proceedings of the Aristotelian Society* (Vol. 104, pp. 53–73). New Series.

Krifka, M. (1989). Nominal reference, temporal constitution and quantification in event semantics. In R. Bartsch, J. van Benthem, & P. van Emde Boas (Eds.), *Semantics and contextual expression* (pp. 75–115). Dordrecht: Foris.

Landman, F. (1989). Groups II. *Linguistics and Philosophy*, *12*(6), 723–744.
Lasersohn, P. (2005). Context dependence, disagreement, and predicates of personal taste. *Linguistics and Philosophy*, *28*(6), 643–686.
Lasersohn, P. (2009). Relative truth, speaker commitment, and control of implicit arguments. *Synthese*, *166*(2), 359–374.
Lassiter, D. (2011). *Measurement and modality: The scalar basis of modal semantics*. Dissertation, New York University.
McNally, L., & Stojanovic, I. (2017). Aesthetic adjectives. In J. O. Young (Ed.), *Semantics of aesthetic judgment* (pp. 17–37). Oxford: Oxford University Press.
Moltmann, F. (2010). Relative truth and the first person. *Philosophical Studies*, *150*(2), 187–220.
Pflaum, T., Konitzer, K., Hofmann, T., & Koehler, P. (2013). Influence of texture on the perception of saltiness in wheat bread. *Journal of Agricultural and Food Chemistry*, *61*(45), 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, K. J. (2009). Judgment ascriptions. *Linguistics and Philosophy*, *32*(4), 327–352.
Sassoon, G. W. (2007). *Vagueness, gradability and typicality, a comprehensive semantic analysis*. Dissertation, Tel Aviv University.
Sassoon, G. W. (2010). Measurement theory in linguistics. *Synthese*, *174*(1), 151–180.
Sassoon, G. W. (2011). Adjectival vs. nominal categorization processes: The rule vs. similarity hypothesis. *Belgium Journal of Linguistics*, *25*, 104–147.
Sassoon, G. W. (2012). The double nature of negative antonymy. In A. Aguilar Guevara, A. Chernilovskaya, & R. Nouwen (Eds.), *Proceedings of Sinn und Bedeutung 16. MIT Working Papers in Linguistics* (pp. 543–556).
Sassoon, G. W. (2013). A typology of multidimensional adjectives. *Journal of Semantics*, *30*(3), 335–380.
Sassoon, G. W. (2015). *A degree-approach account of multidimensional gradability*. Unpublished manuscript, Bar-Ilan University.
Solt, S. (2016). Ordering subjectivity and the absolute/relative distinction. In N. Bade, P. Berezovskaya, & A. Schöller (Eds.), *Proceedings of Sinn und Bedeutung* (Vol. 20, pp. 676–693).
Stephenson, T. (2007). Judge dependence, epistemic modals, and predicates of personal taste. *Linguistics and Philosophy*, *30*(4), 487–525.
Stojanovic, I. (2007). Talking about taste: Disagreement, implicit arguments, and relative truth. *Linguistics and Philosophy*, *30*(6), 691–706.
Suzuki, A. H., Zhong, H., Lee, J., & Martini, S. (2014). Effect of lipid content on saltiness perception: A psychophysical study. *Journal of Sensory Studies*, *29*(6), 404–412.
Umbach, C. (2016). Evaluative propositions and subjective judgments (pp. 127–168). In C. Meier & J. van Wijnbergen-Huitink (Eds.), *Subjective meaning: Alternatives to relativism*. Berlin: de Gruyter.