

Value and Scale: Some Observations and a Proposal

Andrés Soria-Ruiz*

Received: 3 December 2020 / Revised: 7 May 2021 / Accepted: 12 June 2021

Abstract: In this paper, I investigate the scalar semantics of evaluative adjective in general, and of *good* in particular. Lassiter (2017) has argued that *good*, when taking propositions as arguments, has an interval scale. I argue that there's evidence in support of the view that *good*, when taking individuals as argument, has a scale that is stronger than interval, but weaker than ratio. In particular, I propose that individual-level *good* has a “round” ratio scale, which allows a broader set of ratio transformations than standard ratio scales. This conclusion is consistent with the fact that *good* admits round ratio modifiers (*twice as good*), but eschews precise ones (*# 1.38x as good*). An important consequence of this view is that the scales of individual and propositional-level *good* are severed.

Keywords: Evaluative adjectives, scalar semantics; metaethics.

“All my life I'd heard people tell their black boys and black girls to “be twice as good,” which is to say “accept half as much.” These words would be spoken with a veneer of religious nobility, as though they evidenced some unspoken quality, some undetected courage, when in fact all they evidenced was the gun

* Universidade NOVA de Lisboa

 <https://orcid.org/0000-0002-4592-9783>

 Campus de Campolide 1099-085 Lisboa, Portugal.

 aruiz@fcs.unl.pt



to our head and the hand in our pocket.” (Ta-Nehisi Coates, *Between the World and Me*)

1. Introduction

This paper puts forward a puzzle about the semantics of evaluative adjectives, in particular about the adjective *good*. The puzzle is the following: even though *good* largely eschews measurement, phrases like *twice as good* are perfectly interpretable. What do they mean? And what consequences does their acceptability have for the semantics of these adjectives? The purpose of this paper is to investigate those questions.

Evaluatives¹ are gradable predicates, which is attested by the fact that they admit ADJECTIVAL MODIFIERS. To see this, compare (1), where a gradable adjective (*good*) is modified, with (2), where a non-gradable adjective (*hexagonal*) is modified (# indicates that the construction is unacceptable):

- (1) The courtyard is {very} good / {much / a little / better than the park}.
- (2) The courtyard is {# very} hexagonal / {# much / # a little / # more hexagonal than the park}.²

Adjectival modification is a window into the scalar properties of gradable adjectives (Lassiter 2017; Sassoon 2010; Solt 2018, a.m.o.). Different modifiers can tell us different things about the scale corresponding to the relevant adjective. For instance, an adjective like *tall* admits measure phrases, while *good* does not:

- (3) Ann is 180cm tall.
- (4) Bill is ??? good.

¹ This is a heterogeneous class of adjectives whose most eminent members are *good* and *bad*, but which also contains adjectives of moral (*virtuous*), aesthetic (*beautiful*) and personal taste evaluation (*tasty*), as well as so-called thick adjectives (*cruel*).

² This does not mean that the sentences in (2) are absolutely unintelligible; but in order to recover a meaning one needs to do some interpretative work. For instance, *the courtyard is a lot more hexagonal than the park* could mean that it has a more regular and/or carefully delineated hexagonal shape.

Interestingly, one would not even know what to fill the blank with. Bill is... “2 hours good”? “4 agreeable encounters good”? This suggests that, while the scale corresponding to *tall*, i.e., height, admits of measurement, the scale corresponding to *good* does not. There are no standard measures of how good things are.³

Regardless, I want to argue that *good* does not fully eschew measurement. Indeed, my purpose is to show that the scale of *good* poses a puzzle. Among the two most salient scale types used in social sciences, ratio and interval, it is difficult to determine which of these corresponds to *good*. On the one hand, if *good* had an interval scale, it ought to reject ratio modifiers. But individual-level *good* (*x* is *good*) admits round ratio modifiers (viz. Coates’ quote). On the other hand, if *good* had a ratio scale, it ought to be positive with respect to concatenation. This means, roughly, that the combined goodness of any two individuals taken together must be greater than the goodness of each individual taken separately. But according to Lassiter (2017), propositional-level *good* (*it is good that φ*) is not positive with respect to concatenation. Thus, one is confronted with a puzzle. The way out will be to assume that *good* has a different scale when it takes propositions

³ There are exceptions: one can speak of a swimmer being 6 seconds better than another; or a politician doing 3 points better than their opponent on a poll. However, it is intuitive to interpret *better* in those examples as meaning 6 seconds, or 3 points greater on some contextually salient scale, which may not be the scale of *good*. The presence of specific units of measurement in those examples indicates as much: *6 seconds better* suggests that better there just means *faster*; *3 points better* indicates that *better* stands for *greater*. Such reinterpretations can coerce *good* into admitting exact measurements, and moreover, into shifting its scale-type: in both examples, *better* will adopt a ratio scale in virtue of the fact that the relevant properties (speed, score on a poll) have ratio scales. Examples abound: *Apple performed ... 4.5 times better than Blackberry* (<https://www.tradegecko.com/blog/supply-chain-management/apple-the-best-supplychain-in-the-world>). Relatedly, Lassiter (2017, p.89) discusses an Internet example where a company is described as retaining users *2-3 times as efficiently* as another. Independently of whether *efficient* or *good* really are ratio adjectives, in these contexts they behave as such and thereby admit the relevant ratio modifiers. I will not, however, rely on examples like these to conclude that *good* allows exact measurement or has a ratio scale. I thank two reviewers for pressing me to clarify this. In what follows, I set such coerced usages aside.

and when it takes individuals as arguments. In the former case, I will conclude with Lassiter that *good* has an interval scale. In the latter case, however, there are reasons to conclude that *good* doesn't have an interval scale. More specifically, I will propose that individual-level *good* has a scale that is stronger than interval, but weaker than ratio, a scale which I call ROUND RATIO scale, which admits a broader set of ratio transformations than standard ratio scales.

Before moving on, here is a comment on *good*. The hypothesis that *good* has a different scale-type when ranging over individuals and when ranging over propositions is bound to strike as controversial. But this may seem less surprising in light of the fact that the meaning of *good* is massively underspecified: *good* can be interpreted as categorical (*unconditional good*) or hypothetical (*good given certain ends or purposes*); relatedly, *good* has so-called "attributive" and "predicative" uses.⁴ Moreover, *good* is judge-dependent (Bylinina 2017) and multidimensional (Sassoon 2013). Similarly, *good* carries a beneficiary argument – as *in good for you!* (see Stojanovic 2016, pp. 19-20), and even lives a double a life as an intensifier, as in *a good dose of luck* (Castroviejo and Gehrke 2019). Indeed, Hare (1952, see also Umbach 2016) held that the only thing that tied together all uses of *good* was the expression of commendation. In light of such underspecificity, the prospect of assigning different scale types to different uses of *good* may seem less striking.⁵

The paper is laid out as follows: in section 2, the typology of scales standardly used in linguistics is discussed, and the significance of various types of inferences and modifiers is introduced. In sections 3-5, various scale

⁴ The distinction comes from Geach 1956. See e.g., Asher 2011; Ridge 2014; Thomson 2008 for discussion.

⁵ However, this invites a further question. Given that its meaning is so open and minimal, why the focus on *good*? Other adjectives (*beautiful*, *ugly*, *interesting*) may be a bit more uniform, and thus a more reliable guide to the scalar properties of evaluative expressions. Nevertheless, there are reasons to study the semantics of *good*; after all, *good/bad* are the most basic evaluative adjectives. This is evidenced by the fact that all other evaluative adjectives imply *good/bad* in some way on another. Moreover, *good/bad* are some of the few evaluative adjectives to take both individuals and propositions as arguments, which is crucial for my discussion. I thank a reviewer for this journal for pressing me in this regard.

types are introduced (in order of increasing strength) and rejected as candidates for the scale of *good*. In section 3 it is shown that *good* cannot have a merely ordinal scale. Subsequently, I argue that the hypotheses that *good* has an interval (section 4) or ratio scale (section 5) are problematic. In section 6, I propose a solution. Section 7 concludes.

2. Scale structure

Lassiter (2017), Sassoon (2010), and others have resorted to Representational Measurement Theory (RMT, see Krantz et al. 1971) to explore the features of linguistically gradable items. Lassiter in particular focuses on epistemic, probability and deontic modals, but also on the evaluative adjective *good*, while Sassoon 2010 considers a more traditional set of gradable adjectives. In this and the following sections, I present the standard typology of scales following mainly Lassiter, as well as the relevant linguistic tests that can help diagnose the scale type of a scalar item, and we will see what best applies to *good*.

In RMT, the properties of scales are studied by considering what mathematical operations they support. The outcome of this is a typology of scales, or a set of scale types. Lassiter proceeds by attempting to subsume the scales lexicalized in various natural language expressions under scale types defined by RMT. His procedure is roughly the following: starting from the observation that some predicates are gradable, he assumes that they denote scalar properties, or SCALES for short. Then, in order to study the properties of those scales, he does two things. The first is to observe what kind of inferences and modifiers those natural language items allow and forbid. The second is to map the various acceptable uses of those scalar items onto different mathematical relations over the real numbers, in the way that RMT tells us to. Depending on the kind of mapping from natural language onto such mathematical relations that are admissible, a scale can be subsumed under one or other scale type.

For concreteness, let us define a SCALE as a tuple $\mathcal{S} = \langle X, \succcurlyeq, \dots \rangle$ containing a set of individuals X , a binary ordering relation \succcurlyeq and potentially other operations. In order to determine the features of \mathcal{S} , one seeks to define a structure-preserving mapping (a homomorphism) μ from \mathcal{S} onto $\langle \mathbb{R}, \geq, \dots \rangle$

(where \mathbb{R} is the set of real numbers, \geq is the usual ordering relation and other operations over \mathbb{R} might be taken into account). If a function μ is a homomorphism from \mathcal{S} onto $\langle \mathbb{R}, \geq, \dots \rangle$, then μ is called an ADMISSIBLE MEASURE FUNCTION of \mathcal{S} . And to prove that μ is an admissible measure function of \mathcal{S} is to prove a REPRESENTATION THEOREM. Different scale types are then distinguished by imposing different representation theorems that the admissible measure functions must satisfy; the more conditions they must meet, the more structure the scale has – the stronger the scale is.

There is a potentially infinite number of relations that one can define over a scale \mathcal{S} . But the crucial ones for our purposes are the binary ordering relation \succcurlyeq , which was already mentioned, and the operation of CONCATENATION, (which is represented as \circ). Concatenation allows us to construct compound objects from the simple elements in a given domain. For any elements a, b in some domain, $a \circ b$ is the concatenation of a and b .

However, concatenation is not a linguistic operation. In order to represent concatenation in natural language, it has to be mapped onto some model-theoretical relation. Lassiter (2017, p. 39), following Krifka (1989), maps it to the set-theoretical operation of JOIN, \sqcup , restricted to non-overlapping individuals: $x \circ y$ is defined if and only if

1. x and y belong to the same semantic type α , and
2. x and y do not overlap.

When defined, $x \circ y = x \sqcup y$, where \sqcup is JOIN over domain D_α .

JOIN maps onto different aggregation operations depending on what domain one considers. If one considers individuals, JOIN is mereological sum. Thus, for any non-overlapping individuals x, y , $x \sqcup y =$ the complex individual formed by x and y , $x \oplus y$. If one considers propositions, given that the JOIN operation over the domain of propositions amounts to set union, the concatenation of propositions will be their union: for any non-overlapping propositions u, v , $u \sqcup v = u \cup v$, which is represented linguistically as Boolean disjunction. Concatenation is crucial because mapping it to different mathematical relations helps define different scale types (in particular ratio and interval scales).

Lassiter focuses on the three main type of scales used in social and empirical sciences, namely ORDINAL, INTERVAL and RATIO scales. These scales

are defined, as I said above, via their admissible measure functions. In order to investigate what scale a given lexical item has, one needs to consider adjectival modifiers. Adjectival modifiers carry information about the scales of the adjectives that they modify. Among the adjectival modifiers that can offer insight into scale structure there are measure phrases (*two meters, 35 years*), quantity adverbs (*much, a little, a lot*) and ratio modifiers (*twice, 1.38x*).

I will focus on patterns of acceptability and unacceptability. If an adjective accepts a modifier, then I will conclude that the adjective has a scale at least as strong as to represent the information contributed by the modifier. For example, in the introduction it was shown that (3) is an acceptable thing to say. This suggests that the scale of *tall* accepts units of measure based on centimeters (which in turn are based on ratio transformations, cf. Sassoon 2010).

On the other hand, if a modifier is not acceptable, this implies that the scalar information introduced by the modifier is too strong for the relevant adjective. For example, ratio modifiers such as *n-x* (*n-times*) require an adjective with a ratio scale to be interpretable. If they combine with an adjective that has a weaker scale, i.e., *hot, late* or *safe*, it leads to infelicity:

- (5) a. # Bowl A is 1.38x as hot as bowl B.
 b. # Amir came 2x as late as Mora.
 c. # My neighborhood is 4x as safe as yours.

As we will see, *hot* has an interval scale, according to which zero points are variable. *1.38x* requires a fixed zero point, and this is why it cannot combine with *hot*. The scale of *hot* does not provide a fixed zero point that *1.38x* can be interpreted relative to. Nonetheless, making explicit reference to a particular scale (e.g., Celsius) is a repair strategy for sentences like (5a):

- (6) ✓ Bowl A is 1.38x as hot as bowl B on the Celsius scale.

The reason why this qualification is successful is that making reference to the Celsius scale introduces the zero point needed to interpret the ratio 1.38.⁶

⁶ To some speakers, sentences like (5a) sound fine. Erich Rast (p.c.) suggests the following example: *I baked the Beef Wellington twice as hot as Gordon Ramsay said:*

I turn now to presenting these three scale types and to consider whether the scale lexicalized by *good* satisfies each of them.

3. Ordinal scales

A scale \mathcal{S} that is merely ordinal is such that all that can be said of the elements in its domain is how they are ordered with respect to each other. For this reason, all measure functions μ that preserve the ordering among the reals are admissible measure function of \mathcal{S} . No other structure is represented; we do not know anything about the distances between elements on the scale, for instance, or their respective distances to a zero point. The relation of set inclusion is an example of a relation with a merely ordinal structure: all the information that set inclusion represents is an ordering on its domain. More precisely:

Definition 1 (Ordinal scale). *If a scalar property $\mathcal{S} = \langle X, \succcurlyeq \rangle$ is an ORDINAL SCALE (disregarding concatenation and other operations), then every admissible measure function μ that maps \mathcal{S} onto $\langle \mathbb{R}, \geq \rangle$ is such that, for all $x, y \in X$ and $x \succcurlyeq y$, $\mu(x) \geq \mu(y)$.*

Any measure function μ is an admissible measure function of \mathcal{S} as long as, to any two elements x, y of X that stand in the \succcurlyeq relation of \mathcal{S} , μ assigns numerical values such that the value of x is a number at least as great as y . **Definition 1** says nothing about the type of mathematical operation that concatenation should be mapped onto. Thus, any mathematical operation is admissible; it could be addition, subtraction or what have you.

To see how this works, consider again the set inclusion relation. Let us represent it as a structure $\langle \mathcal{P}(X), \supseteq \rangle$, where $\mathcal{P}(X)$ is the power set of some set X , and \supseteq is the superset relation. If this structure is ordinal, then every measure function that maps it onto $\langle \mathbb{R}, \geq \rangle$ should satisfy the representation theorem above. This implies that any mapping that respects the ordering among reals will be an admissible representation of the superset relation.

I set the oven to 200 C instead of 100 C, it's burnt! This deserves more attention, but it's possible that, in many contexts, certain scales are so prevalent that repair strategies such as (6) are not needed.

For any two elements Y and Z of $\mathcal{P}(X)$ such that $Y \supseteq Z$, a μ such that $\mu(Y) = 5$ and $\mu(Z) = 3$ is an admissible measure function; another μ' such that $\mu'(Y) = 12,351$ and $\mu(Z) = -0.0004$ also is; but a μ'' such that $\mu''(Y) = 2$ and $\mu(Z) = 3$ will not do, because $2 \not\geq 3$. The crucial feature of ordinal scales is that nothing matters beyond order; which is why, if one defines scales by its admissible measure functions, ordinal scales are very liberal.

Might the scale of *good* be merely ordinal? Lassiter's answer (2017, p.177), with which I agree, is negative: the scale of *good* cannot be merely ordinal, because in addition to order, the distance between elements also matters.

The crucial data point here is the admissibility of quantificational adverbs such as *much*, *a little* or *a lot*. Note that there is an interpretative difference between the following two sentences:

- (7) Volunteering is better than donating.
- (8) Volunteering is much better than donating.

However vague and variable the meaning of *much* may be, the fact that one can imagine a situation in which (7) would be true while (8) false suggests that they do not mean the same thing.

- (9) Volunteering is better than donating, but not much better.

Informally, the admissibility of such modifiers imposes the requirement on the scale of *good* that the distance between measures be meaningful: (9) says that the value of volunteering is higher than the value of donating, but that the distance between them is not "much". If the scale of *better* were merely ordinal, all measure functions that respect the ordering between the two *relata* should be acceptable. *A fortiori*, measure functions according to which the difference in value between volunteering and donating amounts to "much" and measure functions according to which it does not should all be acceptable. But if that were so, it would not be possible to represent the contrast in truth conditions between (7) and (8). (9) attests such a contrast, and therefore the scale of better cannot be just ordinal.

More formally, the reasoning is the following: if *good* had an ordinal scale, then for any two elements on that scale that are ordered with respect to each other, all measure functions that respect that ordering should be

admissible. A sentence like (9) however, admits certain order-preserving measure functions but also rules out others, namely those that assign a value to each element that is at least as great as whatever quantity *much* stands for:

$$(10) \quad (9) = \mu(\textit{volunteer}) > \mu(\textit{donate}) \ \& \ [\mu(\textit{volunteer}) - \mu(\textit{donate})] \geq / \textit{much}$$

The fact that the truth-conditions of (9) require ruling out certain order-preserving measure functions suggests that the scale of *good* must have more structure than that of an ordinal scale. Regardless of how one defines *much*, there will be order-preserving measure functions for which the relation in (7) holds, but the one in (8) doesn't – just think of any measure function assigning some but not much difference in value to volunteering and donating.

Based on this, one can conclude that *good* must have a stronger scale than ordinal. The reader can check that similar observations apply to other evaluative adjectives, as they can all be modified by quantificational adverbs such as *much*, *a little* or *a lot*. The other two salient alternatives are interval and ratio scales, in order of increasing strength.

4. Interval scales

Interval scales are stronger than ordinal scales, but weaker than ratio scales. They are stronger than ordinal scales because over and above mere order, the distance between elements on the scale, that is, their intervals, matters. However, they are weaker than ratio scales, because they do not determine a zero point, and therefore the positions of elements on the scale cannot be defined using ratios. Interval scales take into account the distance, or gaps, between elements – for this reason, the elements on an interval scale are not actually points, but intervals (although this will not matter for our purposes).

Temperature, clock time or danger are familiar examples of interval scales. Informally, what is crucial about those natural language cases is that the scales that those expressions lexicalize do not determine a zero point: a “zero” degree of temperature is a mere convention, and changes when one

moves from the Celsius to the Fahrenheit scale. Similarly, it is intuitive to think that there is no zero point in clock time or in a scale of danger/safety. Formally, this is cashed out by making ratio transformations meaningful only relative to some arbitrary reference point:

Definition 2 (Interval Scale). *Where $\mathcal{S} = \langle X, \succcurlyeq, \circ \rangle$ is a scale, if \mathcal{S} is an INTERVAL SCALE, then the following representation theorem holds for every admissible measure function μ that maps \mathcal{S} onto $\langle \mathbb{R}, \geq, + \rangle$: for all $x, y \in X$ and $x \succcurlyeq y$,*

(i) $\mu(x) \geq \mu(y)$ and

(ii) for any μ' satisfying condition (i) and for any $z \in X$, there are some n, m such that $n \in \mathbb{R}^+$ and $m \in \mathbb{R}$, $\mu'(z) = n\mu(z) + m$.

That interval scales are strictly stronger than ordinal scales is easily seen by considering that the set of admissible measure functions according to **Definition 2** is a proper subset of the admissible measure functions according to **Definition 1**.

The crucial linguistic prediction associated with interval scales is that ratio modifiers are unacceptable. Recall the following examples:

- (5) a. # Bowl A is 1.38x as hot as bowl B.
 b. # Amir came 2x as late as Mora.
 c. # My neighborhood is 4x as safe as yours.

Those ratio comparisons are meaningless unless a zero point is defined on the relevant scale, but the adjective does not provide one.

So, an attractive explanation for why (5a) is odd becomes available: note that temperature is measured by scales such as Celsius, Fahrenheit or Kelvin. Now, (5a) might be true in a certain scale (say, Celsius). But if one moves to a Fahrenheit scale, the ratio 1.38 will be meaningless because the conversion between Celsius and Fahrenheit does not preserve ratios. For instance, if bowl A is 62.1°C and bowl B is 45°C, one could say that bowl A is 1.38x hotter on the Celsius scale than bowl B. But in a Fahrenheit scale, those temperatures are 143.78 and 113 respectively, and the ratio between them would no longer be 1.38.

However, I noted that (5a) can be repaired by mentioning a specific scale: if one adds the qualification that one is using a Celsius scale, the sentence immediately improves.

(6) ✓ Bowl A is 1.38x as hot as bowl B on the Celsius scale.

The reason for this is that mentioning the Celsius scale introduces the necessary zero point required to interpret the ratio modifier.

Note, in addition, that the comparative size of intervals can be measured using ratios. So even though it does not make sense to say that Amir came twice as late as Mora, it does make sense to say that Amir was delayed by twice as much, or that he stayed for twice as long as Mora. This is because, even though the scale of temporal instants does not have a natural zero point, the intervals between temporal instants do.

Might *good* have an interval scale? The answer is not straightforward. On the one hand, *good* (and evaluatives in general) eschews precise ratio modifiers such as 1.38x.⁷ In this sense *good* behaves like interval adjectives:

(11) # Volunteering is 1.38x as good as donating.

However, this is not enough to conclude that *good* has an interval scale, for two reasons. First, attesting the unacceptability of ratio modifiers is not enough to determine that the relevant adjective has an interval scale. Postulating an interval scale is appropriate in the case of temperature or clock time, but that is because we know independently how temperature and clock time are measured – and, in particular, we know that zero points on the relevant scales are arbitrary. This reasoning does not apply to *good*: we do not know whether putative zero points on the goodness scale are arbitrary, because – again, setting aside coercive interpretations – there is no standard way of measuring value. Given that we lack independent evidence for or against the presence of arbitrary zero points on the *good* scale, we cannot conclude from the unacceptability of precise ratio modifiers that *good* has an interval scale.

Secondly, *good* (and other evaluatives) admit round ratio modifiers, whereas interval adjectives reject them:

(12) You have to be twice as good.⁸

⁷ Recall that uses in which *good* is coerced into a ratio interpretation are set aside, see n.3.

⁸ Adapted from Coates' quote at the start.

- (13) Your daughter is, like, four times more beautiful.⁹
 (14) He'd have to be ten times more charming than Arnold.¹⁰

Given this, the same reasoning regarding quantificational adverbs applies here. If some ratio modifiers are acceptable, this means that there is an interpretative difference between, e.g., *better* and *twice as good*.

- (15) You have to be better than Concha.
 (16) You have to be twice as good as Concha.

If so, then the scale of *good* must be capable of representing this difference. However indeterminate the meaning of *twice as good* may be, the fact that one can imagine a situation in which (15) would be true while (16) is false suggests that they do not mean the same thing.

- (17) You have to be better, but not twice as good as Concha.

As discussed above, if an adjective accepts a modifier, then one can conclude that the adjective has a scale at least as strong as to represent the information contributed by the modifier. The admissibility of ratio modifiers imposes the requirement on the scale of *good* that ratios be meaningful. If the scale of *good* were merely interval, then it would not be possible to represent ratios between degrees. But (17) does represent a ratio between value measures, and therefore the scale of *good* cannot be simply interval. However, Lassiter (2017, 89 and ff) has resisted the view that the acceptability of round ratio modifiers is evidence against adjectives like *good* having an interval scale. In his view, round ratio modifiers are hyperbolic and stand for interval modifiers such as *much* or *a lot*. E.g., *ten times more charming* would be a hyperbolic way of saying *much more charming*. Lassiter says that the fact that those sentences become unacceptable when one adds an adverb like *exactly* points in this direction:

- (18) You have to be (# exactly) twice as good as Concha.
 (19) Your daughter is, like, (# exactly) four times more beautiful.
 (20) He'd have to be (# exactly) ten times more charming than Arnold.

⁹ Adapted from the series *Fresh Off the Boat*, season 5 chapter 5, 2018.

¹⁰ Adapted from the movie *Pulp Fiction*, 1994.

But it is one thing to say that these modifiers are hyperbolic, and a different one to say that they are really interval. The latter view predicts that such modifiers should be admissible with interval adjectives across the board. This prediction is not borne out, as *twice as hot* is just as bad as *1.38x hotter* (the same goes for (5b)-(5c), barring possible acceptable instances, see n.6):

(21) # Bowl A is twice as hot as bowl B.

To avoid this bad prediction, one may reject Lassiter's exact version of a hyperbole view, according to which round ratio modifiers are tantamount to interval modifiers. Alternatively, one could say that modifiers like *10x*, *20x* or *50x* are hyperbolic ways of saying *many times*. This possibility is suggestive when one considers the relative frequency of these modifiers in corpora: briefly, "very" round modifiers such as *2x*, *10x*, *100x* and *1000x* are significantly more frequent than *3x*, *4x*, *5x*, *20x* or *50x*. This suggests that the former might be somewhat idiomatic, and not to be taken as literally expressing measurement.¹¹ But even so, if *good* accepts a modifier like *many times*, this is still evidence that *good* has a stronger scale than interval, contrary to Lassiter.

In sum: partially based on his view that round ratio modifiers are hyperbolic, Lassiter maintains that *good* has an interval scale. But his reasoning is essentially abductive: given that according to him an ordinal and a ratio scale can be ruled out, only interval scales remain as a candidate among the type of scales attested in natural language. I have offered an argument against the view that *good* has an ordinal scale (acceptability of *much*); as well as an argument against the view that *good* has an interval scale (acceptability of *twice*). Lassiter rejects the latter, but I've pushed back against his alternative view that round ratio modifiers are hyperbolic

¹¹ A search on Corpus of Contemporary American English (COCA, <https://www.english-corpora.org/coca/>) of a set of round ratio modifiers between 1 and 100, in addition to 1000, combined with *better* reveals that a handful of round modifiers are significantly more frequent than others. In decreasing frequency: *twice as good* (113), *ten times* (84), *a thousand times* (63), *a hundred times* (38), *five times* (18), *three/four times* (15), *twenty times* (4) and *fifty times better* (3). I thank a reviewer for inviting me to look into this.

interval modifiers. However, I haven't yet looked at Lassiter's argument against *good* having a ratio scale. To this end, let us move on to ratio scales.

5. Ratio scales

Ratio scales are characterized by the fact that the relative “size” of elements matters. In particular, difference in size between elements is measured in ratios, which means that only ordering-preserving measure functions that are obtained via a multiplication operation are admissible. In addition to this, ratio scales require that concatenation be mapped onto the mathematical operation of addition. That is, the concatenation of two elements may only be mapped onto a measure function that assigns to such compound object the arithmetical sum of the individual measures of the concatenated elements.

Scales like height and weight are familiar examples of ratio scales, where the relation between elements in the scale can be mapped onto measure functions that maintain a constant ratio between the numerical values assigned to them. More formally:

Definition 3 (Ratio Scale). *If a scalar property $\mathcal{S} = \langle X, \succsim, \circ \rangle$ is a RATIO SCALE, then the following representation theorem holds for every admissible measure function μ that maps \mathcal{S} onto $\langle \mathbb{R}, \geq, + \rangle$: for all $x, y \in X$ and $x \succsim y$,*

(i) $\mu(x) \geq \mu(y)$,

(ii) $\mu(x \circ y) = \mu(x) + \mu(y)$ and

(iii) for any μ' satisfying (i) and (ii), there's an $n \in \mathbb{R}^+$ s.t. for any $z \in X$, $\mu'(z) = n\mu(z)$.

Recall that admissible measure functions for ordinal scales satisfy only the first of those conditions. Admissible measure functions for interval scales satisfy the first condition as well as a “liberal” version of the third, where ratios are calculated relative to arbitrary and variable “zero” points. Since no reference is made here to such variable, ratios are fixed relative to the real zero. Thus, whereas an interval scale admits all ratios calculated taking any real as reference point, a ratio scale admits only those calculated relative to 0. This means that a ratio scale imposes more conditions on the

admissible measure functions, and is therefore a stronger scale type than ordinal and interval scales.

In order to see how ratio scales constrain admissible measure functions, consider a familiar example: height. Seeing why the height scale \mathcal{S}_{height} is stronger than an ordinal scale is straightforward: suppose that Amir is taller than Mora. If height were an ordinal scale, one should be able to map Amir and Mora's heights to any pair of numerical values under the $>$ relation. But some of those values would radically misrepresent their heights. Suppose that Amir and Mora are respectively 182 and 165 centimeters tall. Consider a measure function μ' that assigns $\mu'(Amir) = 182$, $\mu'(Mora) = 165$, but such that their concatenated heights, $\mu'(Amir \circ Mora)$, is equal to 17. μ' respects the ordering relation between them – i.e., complies with condition (i) in **Definition 3**. But it radically misrepresents the intuitive value of their concatenated heights – it doesn't comply with condition (ii). Or consider another measure function μ'' that assigns $\mu''(Amir) = 182$, $\mu''(Mora) = 181.9$ and $\mu''(Amir \circ Mora) = 363.9$. This measure function respects the ordering relation between Amir and Mora – complying with (i), and the fact that their combined heights should be the arithmetical sum of their individual heights – complying with (ii). But it does not respect the intuitive relation that holds between Amir and Mora's heights, because it does not preserve the ratio between their heights. That is, it does not respect condition (iii) in **Definition 3**.

Condition (ii) and (iii) in **Definition 3** impose more structure on the admissible measure functions for a ratio scale than mere preservation of order, and thereby define a stronger scale. In particular, if a scale \mathcal{S} is ratio, only order-, addition-, and ratio-preserving measure functions are admissible.

I have argued that *good* cannot have an ordinal scale, and I have argued that it doesn't have an interval scale either. Is the goodness scale a ratio scale? There are two considerations against this. First, since ratio scales make ratio comparisons interpretable, adjectives that have a ratio scale are predicted to admit ratio modifiers. This prediction is borne out for *tall*, which is (independently) known to have a ratio scale:

$$(22) \quad \text{Amir is 1.38x as tall as Mora.}$$

Conversely, adjectives that eschew ratio modifiers are predicted to not have ratio scales. Such is the case for *good* and precise ratio modifiers:

- (11) # Volunteering is 1.38x as good as donating.

Secondly, ratio scales are by definition POSITIVE with respect to concatenation, and the scale of *good* is not, according to Lassiter (2017, p. 179 and ff). Being positive with respect to concatenation means that the concatenation of any two elements has a greater degree of the relevant property than either element. More formally, a scale $\mathcal{S} = \langle X, \geq, \circ \rangle$ is positive with respect to concatenation iff for any $x, y \in X$ that do not overlap, $x \circ y > x$ (except if \mathcal{S} is lower-bounded, and y has exactly the value of the lower-bound; i.e., if y is equal to 0). Lassiter holds that the *good* scale lacks this property, based on the observation that it seems to respect the following inference pattern:

- (23) a. $a \geq b$
 b. $a \geq c$
 $\therefore a \geq (b \circ c)$

If \mathcal{S} were positive with respect to concatenation, that inference should fail in many instances. But it does not fail for *good* (by contrast, it very clearly fails for *likely*, which is independently argued to have a ratio scale). For an example, consider the following, intuitively valid inference from Lassiter (2017, p. 179; recall that concatenation for propositions is disjunction):

- (24) a. It's as good for the card to be a spade as it is for it to be a heart.
 b. It's as good for the card to be a spade as it is for it to be a diamond.
 \therefore It's as good for the card to be a spade as it is for it to be a red card.

According to Lassiter, that this inference pattern is in general valid shows that the scale of *good* has to be weaker than a ratio scale.

Here appears a dilemma. The first horn is that considerations about precise ratio modifiers suggest rejecting all scale types as weak as ratio. The second horn is that considerations about concatenation suggest rejecting any scale type as strong as ratio, since all are positive with respect to

concatenation. I propose to solve this dilemma by partially rejecting the second horn: Lassiter's considerations about concatenation apply to propositional-level *good* (*it is good that p*), but not to individual-level, or adnominal *good* (*x is good*). Thus, even though there's reason to reject any scale type as strong as ratio for propositional *good*, those considerations do not extend to individual-level *good*. In sum, I agree with Lassiter that propositional-level *good* has an interval scale, but I'll propose that individual-level *good* has a scale that is stronger than interval, although weaker than a standard ratio scale.

When one moves from propositional to adnominal *good*, the inference in (24) arguably fails. Intuitively, this is the case because concatenation for individuals is mereological sum, and the sum of two individuals can have a higher value than each of those individuals taken separately:¹²

- (25) a. Car *a* is at least as good as car *b*.
 b. Car *a* is at least as good as car *c*.
 \therefore Car *a* is at least as good as car $b \oplus c$.

To see how this inference can fail, one can think of good in terms of preference: for any two individuals *x, y*, *x* is at least as good as *y* just in case *x* is at least as preferable as *y*. In turn, one may spell this out by saying that *x* is at least as preferable as *y* just in case every time you have the option of choosing *y*, you also choose *x*. Understood in this way, premise (25a) says that every time you have the option of choosing *b*, you also choose *a*. Premise (25b) says that every time you have the option of choosing *c*, you also choose *a*. But it is consistent with this that if you get to choose the sum of *b* and *c*, you may no longer choose *a* as well. In other words, cars *b* and *c* may have a higher value taken together than taken separately, making the premises true but the conclusion false.

The key to the contrast between adnominal and propositional-level *good* is, of course, concatenation: concatenation for individuals is mereological sum, while for propositions it is disjunction. This has completely different

¹² One might disagree about specific cases, perhaps with other evaluative adjectives. Can the sum of two dishes be tastier than the tastiest of them? Can the sum of two pictures be more beautiful than the most beautiful of the two? Perhaps not, but a single positive instance suffices to falsify the inference, and it can be found.

consequences for the assessment of complex objects. Note that, intuitively, the value of a proposition amounts to the value of its outcome. Similarly, the value of a disjunction must also amount to the value of its outcome, that is, one of its disjuncts. This suggests that the value of a disjunction is maximal, that is, a disjunction is no more valuable than its most valuable disjunct. By contrast, the value of a mereological sum of individuals is potentially positive, that is, higher than the value of each individual in it.

Lassiter relies on examples like (24) to conclude that the scale of propositional *good* can't be ratio, and I agree with his conclusion. But (25) shows that this does not extend to individual *good*. This suggests that propositional and individual good might have different scales. In the next section I will argue that a further observation supports this conclusion, and I will propose a stronger scale type for individual-level *good* than for propositional-level *good*.

6. Round ratio scales

In this last section, I want to propose that individual-level *good* has a stronger scale than interval, but not as strong as a standard ratio scale (as defined in **Definition 3**). Informally, the idea is the following: whereas a standard ratio scale requires that admissible measure functions preserve a precise ratio, which is a positive real, one can define a type of ratio scale according to which this requirement is relaxed, so that admissible measure functions preserve only an approximate ratio. I will call this type of scale a ROUND RATIO SCALE. In practice, this means that a round ratio scale rules out less measure functions than a standard ratio scale, and is thereby weaker.

To define such a scale, one can impose the requirement that the ratio that gets preserved across measures of the same individual is not a positive real, but some positive in its vicinity, defined by a HALO. A halo is an interval around a number whose size can vary. For example, the halo of 2 could be the interval $[1.9, 2.1]$, or $[1.8, 2.2]$.¹³

¹³ See Lasersohn 1999 on how halos are at play in the interpretation of numerical expressions, as well as Hoek 2018; Sauerland and Stateva 2011 for elaboration and criticism of Lasersohn's seminal view.

It is well-known that rounder numbers have greater halos; e.g., 10 has a greater halo than 11 or 9; 50 has a greater halo than 49 or 51.¹⁴ But what numbers are round, and why do they have greater halos? Defining round numbers is not as straightforward as it may seem; for my purposes, I will rely on the following informal and comparative definition: an integer is round just in case it has a larger number of smaller factors than its neighboring numbers (Hardy 1940, p. 48). For example, 10 has more and smaller factors than 9 and 11; the same goes for 50 as opposed to 49 and 51. Even more informally, one tends to consider rounder numbers that end on one or more zeros (relative to a given base), as well as simple multiples or fractions of such numbers (Sigurd 1988, p. 249).

Regarding the question of why rounder numbers have greater halos, one possible answer is that round numbers are cognitively significant (see Rosch 1975, who characterizes round numbers as a kind of *cognitive reference points*). Alternatively, or perhaps as a result of their cognitive significance, round numbers tend to be linguistically simpler (Lotz 1955, see also Krifka 2002, 2007). The cognitive significance of round numbers, sometimes called the *round number bias*, has been studied in domains such as psychology or economics (Lacetera et al. 2012; Lynn et al. 2013; Pope and Simonsohn 2011).¹⁵

Having characterized halos and round numbers, let's now define a round ratio scale. The definition is similar to the standard ratio scale, except that condition (iii) in **Definition 3** above is relaxed, so that admissible ratio transformations are restricted, not to those that preserve some real, but to those that preserve some real within some other real's halo. This is achieved

¹⁴ This is reflected in loose talk. To take an example from Hoek (2018, p. 175), at 3:58 it is preferable to say *4 o' clock* than *3:57*, even though the latter is closer to the truth.

¹⁵ A salient manifestation of the round number bias is the *left-digit effect*, which explains the tendency to price items right below round numbers, such as 3.99€. Buyers perceive the difference between 3.99 and 4 as more meaningful than the difference between, e.g., 4 and 4.01, and sellers take advantage of it (see Bhattacharya et al. 2012). Another interesting manifestation of round number bias is the strive, in sports and other domains, to attain round scores (Lotz 1955; Pope and Simonsohn 2011).

by substituting a fixed ratio for a *halo function*. A halo function H_k is a function from \mathbb{R}^+ to \mathbb{R}^+ such that, for any $m, n, k \in \mathbb{R}^+$, $H_k(n) = m$ just in case m is the result of some simple arithmetical operation on n that maps n onto some number that is no further from n than the halo size of k . For example, suppose that $k = 2$. Assuming that the halo of 2 is the interval $[1.9 - 2.1]$, whose size is 0.2, there's infinitely many functions H_2 , all those functions that take their argument n to any number no further away from n than 0.2. Here are some examples of possible functions H_2 (note that this includes the identity function):

- $H_2^i = \lambda n. n$
- $H_2^j = \lambda n. n + 0.1$
- $H_2^k = \lambda n. n - 0.05$
- ...

But note that, e.g., a function $\lambda n. n + 0.3$ is not such a function, as it maps its argument further away from the halo size of 2.

If, instead of imposing the requirement that admissible transformations preserve some ratio n , one imposes the requirement that they preserve a ratio that results from mapping n to some number in its vicinity, one can allow the necessary variability. Let us define Round Ratio Scales as follows:

Definition 4 (Round Ratio Scale). *If a scalar property $\mathcal{S} = \langle X, \geq, \circ \rangle$ is a ROUND RATIO SCALE, then the following representation theorem holds for every admissible measure function μ that maps \mathcal{S} onto $\langle \mathbb{R}, \geq, + \rangle$: for all $x, y \in X$ and $x \geq y$,*

(i) $\mu(x) \geq \mu(y)$,

(ii) $\mu(x \circ y) = \mu(x) + \mu(y)$ and

(iii) for any μ' satisfying (i) and (ii), there are $n, m \in \mathbb{R}^+$ s.t., for any $z \in X$, there's some function H_m s.t. $\mu'(z) = H_m(n)\mu(z)$.

A round ratio scale does not require that the ratios between the measures assigned to each individual are held constant across admissible measure functions; rather, such ratios are allowed to vary within a certain halo. Thus, for instance, given an admissible measure function μ such that $\mu(x) = 2$ and $\mu(y) = 1$ (for any $x, y \in X$ such that $x \geq y$), consider a measure func-

tion μ' such that $\mu'(x) = 4.1$ and $\mu'(y) = 2$. Given μ , μ' would be inadmissible in a standard ratio scale, since there is no positive real n such that, for every $z \in X$, $\mu'(z) = n\mu(z)$. For x , $\mu'(x)/\mu(x)$ is 2.05; while for y , $\mu'(y)/\mu(y)$ is 2. Those ratios are not the same, and *a fortiori* there does not exist a single ratio for all measure functions applied across all elements of X .

But μ' is an admissible measure function according to **Definition 4**. The reason is that, even though $\mu'(x)/\mu(x) \neq \mu'(y)/\mu(y)$, that is, $2.05 \neq 2$, there exist halo functions H_m , for some positive real m , such that one can map either of these ratios, $\mu'(x)/\mu(x)$ or $\mu'(y)/\mu(y)$, to match some positive real, namely 2. First, consider x . There is a function H_m such that $\mu'(x) = H_m(2)\mu(x)$. This function, call it H_m^i , is $\lambda n.n + 0.05$. Substituting H_m for H_m^i in $\mu'(x) = H_m(2)\mu(x)$, we obtain $4.1 = \lambda n.n + 0.05(2) \times 2$, that is, $4.1 = 2 + 0.05 \times 2$, which is true. Secondly, consider y . There is also a function such that $\mu'(y) = H_m(2)\mu(y)$. This function, call it H_m^j , is just the identity function, $\lambda n.n$. Substituting H_m for H_m^j in $\mu'(y) = H_m(2)\mu(y)$, we obtain $2 = \lambda n.n(2) \times 1$, that is, $2 = 2 \times 1$, which is true.

This opens up the possibility that ratios are calculated only approximately. However, here appears a hurdle: given that the set of reals \mathbb{R}^+ is countably infinite, halos can be of countably infinite size as well. This means that **Definition 4** does not, after all, rule out any measure function (beyond those that fail conditions (i) or (ii)): however different the ratio assigned by two measure functions to a pair of individuals may be, their difference will fall within the halo of *some* real. For example, suppose again that an admissible measure function μ is such that $\mu(x) = 2$ and $\mu(y) = 1$. According to **Definition 4**, any other admissible measure function μ' must be such that there exist $n, m \in \mathbb{R}^+$ such that, for every $z \in X$, there exists some function H_m such that $\mu'(z) = H_m(n)\mu(z)$. The issue is that there will always be some positive real m whose halo is as great as required, so there is in fact no restriction on how far ratios can come apart. In sum, **Definition 4** above is too weak.

Regardless, one can adopt **Definition 4** as a template, and use it to define different round ratio scales of specific granularity. By assigning a specific granularity, one determines a maximum halo size that ratio transformations are allowed to vary within, thereby restricting the admissible measure functions in a way that strengthens Definition 4:

Definition 5 (Round Ratio Scale of n -granularity). *If a scalar property $\mathcal{S} = \langle X, \succcurlyeq, \circ \rangle$ is a ROUND RATIO SCALE OF n -GRANULARITY, then the following representation theorem holds for every admissible measure function μ that maps \mathcal{S} onto $\langle \mathbb{R}, \geq, + \rangle$: for all $x, y \in X$ and $x \succcurlyeq y$,*

(i) $\mu(x) \geq \mu(y)$,

(ii) $\mu(x \circ y) = \mu(x) + \mu(y)$ and

(iii) *for any μ' satisfying (i) and (ii), there's an $m \in \mathbb{R}^+$ s.t., for any $z \in X$, there's some function H_n s.t. $\mu'(z) = H_n(m)\mu(z)$.*

Thus, for example, if a scalar property has a round ratio scale of 2-granularity, then the absolute difference between the ratios assigned by any two admissible measure functions to any pair of individuals cannot be greater than $H(2)$, that is, $[1.9 - 2.1] = 0.2$. Recall μ and μ' . Given the measures assigned to $x, y \in X$, μ' was an admissible measure function according to **Definition 4**. But what about according to **Definition 5**? It depends on whether, for every $z \in X$, there's a function H_2 – a function that maps its argument to a number no further away from it than 0.2 – such that $\mu'(z) = H_2(m)\mu(z)$, for some positive real m . Such functions were already found for x and y , namely $H_m^i = \lambda n.n + 0.05$ and $H_m^j = \lambda n.n$, respectively. Therefore, given μ, μ' is an admissible function for a round ratio scale of granularity 2.

But consider, by contrast, another measure function μ'' according to which $\mu''(x) = 6$ and $\mu''(y) = 2$. Is there some positive real m such that, for every $z \in X$, there is some function H_2 such that $\mu''(z) = H_2(m)\mu(z)$? Recall that it's not necessary for $\mu''(x)/\mu(x)$ to be identical to $\mu''(y)/\mu(y)$, as would be the case in a standard ratio scale. Rather, what is needed is for those ratios to vary at most by 0.2. In other words, the absolute difference between $\mu''(x)/\mu(x)$ and $\mu''(y)/\mu(y)$ has to not be greater than 0.2. But since $\mu''(x)/\mu(x) = 3$ and $\mu''(y)/\mu(y) = 2$, this is not the case. Thus, given μ, μ'' is not an admissible measure function of a round ratio scale of granularity 2.

Importantly, recall that rounder numbers have greater halos. Therefore, the “rounder” its granularity, the weaker a round ratio scale will be: a round ratio scale of 20-granularity will be weaker than one of 2-granularity, which will be weaker than one of 0.2-granularity, and so on. This means that a measure function that is not admissible on a given round ratio scale might

be admissible on a round ratio scale with rounder granularity. Suppose that the halo of 5 is $[4.5 - 5.5]$, that is, 1. As just shown, given μ , μ'' above would not be an admissible measure function of a round ratio scale of granularity 2. But μ'' would be an admissible measure function of a round ratio scale of granularity 5. For that, the absolute difference between $\mu''(x)/\mu(x)$ and $\mu''(y)/\mu(y)$ need not be greater than the halo size of 5, that is, 1. And it isn't. Therefore, bigger halos make for weaker scales.

The proposal is, then, that adnominal *good* has a round ratio scale of n -granularity. But what n ? Settling this requires saying something about ratio modifiers. A simple way of cashing out the meaning of any ratio modifier is to assign to it the presupposition that the adjective with which it combines has a ratio scale, and then assign to it the at-issue meaning one would expect:

- (26) $[[x \text{ is } n\text{-}x \text{ as } A \text{ as } y]] = \text{defined only if } A \text{ has a ratio scale.}$
 If so, $[[x \text{ is } n\text{-}x \text{ as } A \text{ as } y]] = 1 \text{ iff } \mu_A(x) = n\mu_A(y)$

According to this simple proposal, ratio modifiers should be acceptable across the board with ratio adjectives. This prediction is borne out for standard ratio adjectives such as *tall*, as they admit any ratio modifier. But it fails for evaluative adjectives such as *good*, which admit some ratio modifiers (2x) but not others (1.38x).

However, if one modifies the presupposition of ratio modifiers, so that their number indicates the granularity of the ratio scale that they require, their acceptability can serve as a guide to the granularity of ratio adjectives. Here is a proposal:

- (27) $[[x \text{ is } n\text{-}x \text{ as } A \text{ as } y]] = \text{defined only if } A \text{ has a ratio scale of } n\text{-granularity.}$ If so, $[[x \text{ is } n\text{-}x \text{ as } A \text{ as } y]] = 1 \text{ iff } \mu_A(x) = n\mu_A(y)$

According to this view, ratio adjectives have the granularity of the most precise ratio modifier that they accept. Standard ratio adjectives, like *tall*, have maximal granularity, and thus accept any ratio modifier. But other adjectives have less-than-maximal granularity. Adnominal *good*, for example, will be of 2-granularity, since this is likely the most precise ratio modifier that it can admit. More precise ratio modifiers, such as *1.38x*, will require a ratio scale of *1.38-granularity*, which is too precise for *good*, and this is why a phrase like *1.38x better* is infelicitous. In turn, since according

to (27) ratio modifiers presuppose that their adjective has at least their granularity, *m-x as A* will be infelicitous if *A* does not have a ratio scale of at least *m*-granularity.

Moreover, I argued that adnominal *good* is positive with respect to concatenation (cf. (25)). Round ratio scales are so as well, so this prediction is also borne out. In sum, I am claiming that *good*, when taking individuals as its arguments, has a round ratio scale, which is a type of scale that is weaker than standard ratio insofar as it admits more measure functions than standard ratio scales.

Before moving on, an important question remains: why would evaluative adjectives in general, and adnominal *good* in particular, have scales that blur precise ratios? Relying on broad views about vagueness, one may distinguish three genres of response: metaphysical – because there’s no such thing as precise ratios of value, epistemic – because we cannot know objectively where precise ratios are, and psychological – because we are psychologically insensitive to them. Setting aside a metaphysical view, which would require much more discussion than I have space for, an intuitive justification for going in for an epistemic view would be, perhaps, that we simply haven’t yet figured out how to measure value precisely enough in an intersubjectively verifiable way. That is, it may be only subjectively possible to distinguish *2x as good* from *1.9x* or *2.1x* as good. Intersubjective measures of value might just be approximate. This view might be bolstered by the fact, pointed out in the Introduction, that *good* is a judge-dependent and/or multidimensional predicate. The idea could be that, even though each of us may be able to subjectively determine a standard ratio scale of value – perhaps through some operation of dimension aggregation – the best we can do to share such measurements with others are rough approximations, that nevertheless succeed in preserving the overall scalar architecture.¹⁶

A psychological view, on the other hand, might be supported by features of our perceptual and cognitive system. It is well-known that, even though we are capable of representing magnitudes in a mathematically precise way, our perceptual system represents magnitudes in an analog fashion, assigning

¹⁶ I thank two reviewers for this journal for independently pointing to this hypothesis. See also Sassoon 2010, p. 161 and ff., for ideas in this vicinity.

measures to objects that, even though globally covariant with the represented magnitude, introduce a great deal of probabilistic error.¹⁷ It is not wholly counterintuitive to think that, even though the linguistic expression of many such magnitudes (height, weight, distance) inherit their scalar properties from our mathematical capacity to represent precise magnitudes, expressions that denote the *value* of such magnitudes, that is, evaluative adjectives, inherit their scalar properties from our imprecise perceptual system.

Discussing these hypotheses further will have to wait for another occasion. For now, I take it to be at least a competing hypothesis that adnominal *good* has a round ratio scale. By contrast, I conclude with Lassiter that propositional-level *good* has an interval scale. To further support this, note the following: whereas a phrase such as *twice as good* is fairly common, it is hardly ever used to compare the value of propositions. Examples do not abound,¹⁸ and when they are felicitous, they seem to inherit their acceptability from an individual-level comparison. For example, one can say something like:

- (28) It is good that Camila came to the party. It would have been twice as good if she had come with Milica!

Even though this sentence compares the value of propositions, the aggregated value of Camila and Milica coming to the party cannot result from the concatenation of the proposition that Camila comes to the party and the proposition that Milica comes to the party, since the concatenation of those two propositions is their disjunction. Somehow, the aggregated value of Camila and Milica coming to the party is vicariously calculated by aggregating their value as individuals.

Assuming that these examples are rare, and derive their meaning from an individual-level evaluation, the observation that ratio modifiers are unacceptable with propositional-level *good* is predicted if propositional-level *good* has an interval scale, since *twice* is too strong for propositional-level *good*.

¹⁷ This observation is familiar from the literature on vagueness, see e.g., Égré 2017; Fults 2011.

¹⁸ None of the 113 hits of *twice as good* at COCA apply to propositions.

7. Conclusion

In this paper, I have discussed the scale of evaluative adjectives, and, in particular, of the evaluative adjective *good*. I've argued, first, that the lack of measure phrases gives linguistic support to the (otherwise natural) view that *good* lacks measurement units. However, *good* does not eschew measurement altogether. I have argued that there are strong reasons to think that the scale of *good* is stronger than a mere ordinal scale; and I have moved on to discuss which of the two most salient candidates discussed in the literature is most appropriate for *good*, an interval or a ratio scale.

Interval scales were rejected based chiefly on the observation that evaluatives admit round ratio modifiers, but this conclusion was not free of controversy, as Lassiter has argued that those are hyperbolic uses. Ratio scales, by contrast, were partially rejected based on features of concatenation. Ratio scales are by definition positive with respect to concatenation, and it was observed that, while there is no evidence of propositional-level *good* being positive with respect to concatenation, one can argue that individual-level *good* is. Based on this, I proposed to sever the scales of propositional- and individual-level *good*: while the former has an interval scale, the latter has a ratio scale. More specifically, I've proposed that individual-level *good* has a ROUND RATIO SCALE, a type of scale that preserves approximate rather than precise ratios, and is thereby stronger than interval, but weaker than a standard ratio scale.

One potentially controversial aspect of this view remains to be discussed. What are the consequences of severing the scale of propositional- and individual-level *good*? As mentioned in the Introduction, this proposal is bound to be met with resistance, since there's *prima facie* reasons to maintain a uniform view about the scalar semantics of *good*. However, in light of the massive underspecificity of *good*, the prospects of this schism may seem less controversial. In fact, the view that "there's more than one *good*", that is, that *good* is ambiguous or polysemous, is perhaps not such a revisionary hypothesis in light of the properties of *good* reviewed in the Introduction. Indeed, the arguments put forward in this paper might be seen as supporting that general hypothesis. Moreover, the relationship between individual- and propositional-level *good* is understudied, so the view that these

evaluatives might have different scales is not at odds with any existing proposal that I know of. This divergence may simply be one more among other puzzling properties of evaluative adjectives.

Acknowledgments

Special thanks to Erich Rast for multiple comments on various versions of this work. Thanks to Federico Faroldi, Mora Maldonado, Isidora Stojanovic, as well as to two anonymous reviewers for *Organon F* and Dan Zeman for his careful editorial work. This work has been funded by Portuguese national funds through FCT – Fundação para a Ciência e a Tecnologia, I.P., within project UIDB/00183/2020 and by COST Action CA17132, funded by the Horizon 2020 Framework Programme of the European Union.

References

- Asher, Nicholas. 2011. *Lexical Meaning in Context: A Web of Words*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511793936>
- Bhattacharya, Utpal, Craig W. Holden, and Stacey Jacobsen. 2012. “Penny Wise, Dollar Foolish: Buy–sell Imbalances on and Around Round Numbers.” *Management Science* 58 (2): 413–31. <https://doi.org/10.1287/mnsc.1110.1364>
- Bylinina, Lisa. 2017. “Judge-dependence in Degree Constructions.” *Journal of Semantics* 34 (2): 291–331. <https://doi.org/10.1093/jos/ffw011>
- Castroviejo, Elena, and Berit Gehrke. 2019. “Intensification and Secondary Content: A Case Study of Catalan *Good*.” In *Secondary Content: The Semantics and Pragmatics of Side Issues*, edited by Daniel Gutzmann, and Katharina Turgay, 108–43. Brill. https://doi.org/10.1163/9789004393127_006
- Égré, Paul. 2017. “Vague Judgment: A Probabilistic Account.” *Synthese* 194 (10): 3837–65. <https://doi.org/10.1007/s11229-016-1092-2>
- Fults, Scott. 2011. “Vagueness and Scales.” In *Vagueness and Language Use*, edited by Paul Égré, and Nathan Klinedinst, 25–50. Springer.
- Geach, Peter T. 1956. “Good and Evil.” *Analysis* 17 (2): 33–42. <https://doi.org/10.2307/3326442>
- Hardy, Godfrey Harold. 1940. *Ramanujan: Twelve Lectures on Subjects Suggested by His Life and Work*. Cambridge University Press.
- Hare, Richard M. 1952. *The Language of Morals*. Oxford University Press. <https://doi.org/10.1093/0198810776.001.0001>
- Hoek, Daniel. 2018. “Conversational Exculpature.” *Philosophical Review* 127 (2): 151–96. <https://doi.org/10.1215/00318108-4326594>

- Krantz, David H, Patrick Suppes, and Robert Duncan Luce. 1971. *Foundations of Measurement*. New York Academic Press.
- Krifka, Manfred. 1989. "Nominal Reference, Temporal Constitution and Quantification in Event Semantics." In *Semantics and Contextual Expression*, edited by Renate Bartsch, Johan Van Benthem, and Peter Van Ende Boas, 75–115. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110877335-005>
- Krifka, Manfred. 2002. "Be Brief and Vague! And How Bidirectional Optimality Theory Allows for Verbosity and Precision." In *Sounds and Systems: Studies in Structure and Change. A Festschrift for Theo Vennemann*, edited by David Restle, and Dietmar Zaefferer, 439–58. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110894653.439>
- Krifka, Manfred. 2007. *Approximate Interpretation of Number Words*. Manuscript, Humboldt-Universität zu Berlin, Philosophische Fakultät II. <https://doi.org/10.18452/9508>
- Lacetera, Nicola, Devin G. Pope, and Justin R. Sydnor. 2012. "Heuristic Thinking and Limited Attention in the Car Market." *American Economic Review* 102 (5): 2206–36. <https://doi.org/10.1257/aer.102.5.2206>
- Lasersohn, Peter. 1999. "Pragmatic Lalos." *Language* 75 (3): 522–51. <https://doi.org/10.2307/417059>
- Lassiter, Daniel. 2017. *Graded Modality: Qualitative and Quantitative Perspectives*. Oxford: Oxford University Press.
- Lotz, John. 1955. "On Language and Culture." *International Journal of American Linguistics* 21 (2): 187–9. <https://doi.org/10.1086/464329>
- Lynn, Michael, Sean Masaki Flynn, and Chelsea Helion. 2013. "Do Consumers Prefer Round Prices? Evidence from Pay-What-You-Want Decisions and Self-Pumped Gasoline Purchases." *Journal of Economic Psychology* (36): 96–102. <https://doi.org/10.1016/j.joep.2013.01.010>
- Pope, Devin & Uri Simonsohn. 2011. "Round Numbers as Goals: Evidence From Baseball, SAT Takers, and the Lab." *Psychological Science* 22 (1): 71–9. <https://doi.org/10.1177%2F0956797610391098>
- Ridge, Michael. 2014. *Impassioned Belief*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199682669.001.0001>
- Rosch, Eleanor. 1975. "Cognitive Reference Points." *Cognitive Psychology* 7 (4): 532–47. [https://doi.org/10.1016/0010-0285\(75\)90021-3](https://doi.org/10.1016/0010-0285(75)90021-3)
- Sassoon, Galit W. 2013. "A Typology of Multidimensional Adjectives." *Journal of Semantics* 30 (3): 335–80. <https://doi.org/10.1093/jos/ffs012>
- Sassoon, Galit W. 2010. "Measurement Theory in Linguistics." *Synthese* 174 (1): 151–80. <https://doi.org/10.1007/s11229-009-9687-5>

-
- Sauerland, Uli, and Penka Stateva. 2011. "Two Types of Vagueness." In *Vagueness and Language Use*, edited by Paul Egré, and Nathan Klinedinst, 121–45. Springer.
- Sigurd, Bengt. 1988. "Round Numbers." *Language in Society* 17 (2): 243–52. <https://doi.org/10.1017/S0047404500012781>
- Solt, Stephanie. 2018. "Multidimensionality, Subjectivity and Scales: Experimental Evidence." In *The Semantics of Gradability, Vagueness, and Scale Structure*, edited by Elena Castroviejo, Louise McNally, and Galit W. Sassoon, 59–91. Springer. https://doi.org/10.1007/978-3-319-77791-7_3
- Stojanovic, Isidora. 2016. "Expressing Aesthetic Judgments in Context." *Inquiry* 59 (6): 1–23. <https://doi.org/10.1080/0020174X.2016.1208922>
- Thomson, Judith. 2008. *Normativity*. Open Court.
- Umbach, Carla. 2016. "Evaluative Propositions and Subjective Judgments." In *Subjective Meaning: Alternatives to Relativism*, edited by Cécile Meier, and Janneke van Wijnberger- Huitink, 127–68. Berlin: De Gruyter. <https://doi.org/10.1515/9783110402001-008>