Measurement, and interadjective comparisons

Robert van Rooij*
r.a.m.vanrooij@uva.nl
Institute for Logic, Language and Information,
Universiteit van Amsterdam

Abstract

This paper shows the relevance of measurement theory for the linguistic analysis of comparative statements. In particular, the paper focusses on interadjective comparatives like 'x is P-er than y is Q' and comparatives involving multidimensional adjectives. It is argued that Bale's (2008) recent proposal to account for such comparatives is rather limited, and just one way to account for interadjective comparison. In fact, it is shown that we can make use of recently developed measurement-theoretic techniques in political economy to handle intersubjective comparisons of utility to account for interadjective comparatives as well. This paper also discusses how to *construct* the desired scales, if one starts with a delineation approach of comparatives.

1 Introduction

It is sometimes believed that we have two competing analyses of comparatives: one that crucially uses degrees, and another that denies their existence. But it is, of course, absolutely non-sensical to adopt the latter position. Nobody can deny that we sometimes use measure-phrases in natural language, and that they are used in comparatives as well. The issue is rather whether we *always* use degrees when we evaluate comparatives, or whether this is not the case. Moreover, if degrees are essential for the interpretation of a particular

*Parts of this paper were presented during the spring of 2009 in Osnabrueck, Paris, and Amsterdam. I would like to thank the audiences for comments and discussion, Galit Sassoon in particular. This paper is a much improved version of earlier manuscripts, both in content and especially in presentation. I would like to thank the anonymous reviewers of the Journal of Semantics and the editor Bart Geurts for many useful comments and suggestions. I would like to thank Jason Quinley for correcting the English.

kind of comparative, we would like to know *how fine-grained* we should assume these degree-structures to be. Measurement theory provides standard answers to those questions, and it is only natural to assume that measurement theory is relevant for natural language semantics as well (cf. Klein, 1991, Sassoon, 2010).

Measurement theory also helps us to see what is required to account of comparatives like the following:

(1) x is P-er than y is Q.

In linguistics it is generally assumed that to make sense of such comparatives we *must* make use of degrees. Either because both the P-er and the Q-er than relation have the same zero-point and the same unit of measurement (in direct comparisons like 'this table is longer than that table is wide'), or in so-called *indirect*, or *relative* comparisons (Bale, 2008) where the rank of the two items is compared in their respective 'P-er than' and 'Q-er than' orderings. I will show in this paper that Bale's method is rather limited, and just one way to account for interadjective comparison. Other ways are less committing. I will also show that the issue of how to make sense of sentences like (1) is in a sense already adequately handled in measurement theory. What is less clear is how to *construct* the desired scales. Finally, I will suggest that making sense of (1) invites several types of analyses of comparisons involving multidimensional adjectives that linguists normally don't take into serious account.

2 Orders, choice structures, and measurement

Comparative constructions are normally based on gradable adjectives (like *tall* and *warm*). In semantics it is standard to distinguish two types of approaches to the analysis of gradable adjectives: degree-based approaches and delineation approaches. Degree-based approaches (e.g. von Stechow, 1984; Kennedy, 1999) analyze gradable adjectives as relations between individuals and degrees, where these degrees are thought of as scales associated with the dimension referred to by the adjective. Individuals can possess a property to a certain measurable degree. The truth conditions of sentences involving these adjectives are stated in terms of degrees in well-known ways. Delineation approaches analyze gradable adjectives like 'tall' as simple predicates, but assume that the extension of these terms are crucially context dependent. For our purposes, Klein's (1980, 1991) approach is most relevant. Klein (1980) assumes that every adjective should be interpreted with respect to a *comparison class*, i.e. a set of individuals. The truth of a sentence like *John is tall* depends on the contextually given comparison class: it is true in context (or comparison class) c iff John is counted as tall in this class. Klein (1980) further proposes that the meaning of the comparative *John is taller than Mary* is context independent and the sentence is true iff there is a context (henceforth we will use 'context' instead of the more cumbersome 'comparison class') according to which John counts as tall, while Mary does not. If there is any context in which this is the case, it will also be the case in the context containing only John and Mary.

Klein (1980) favors the delineation approach towards comparatives for a number of reasons. First, a degree-based approach only makes sense in case the comparative gives rise to a *total* ordering. But for at least some cases (e.g. *cleverer than*) this doesn't seem to be the case. Second, the delineation account assumes that the meaning of the comparative 'taller than' is a function of the meaning of 'tall', which — according to Klein— is more in line with Frege's principle of compositionality,[1] and also accounts for the fact that in a wide variety of languages the positive is formally unmarked in relation to the comparative.

Van Benthem (1982) has shown how the 'taller than' relation can be *derived* from how we positively use the adjective in certain contexts, plus some additional constraints on how the meaning of the adjective can change from context to context. This is done in terms of the notion of a *context structure*, M, being a triple $\langle X, C, V\rangle$, where X is a non-empty set of individuals, the set of contexts, C, consists of all finite subsets of X, and the valuation V assigns to each context $c \in C$ and each property T those individuals in c which are to count as 'being T in c'.

This definition leaves room for the most diverse behavior of individuals across contexts. Based on intuition (for instance by visualizing sticks of various lengths), however, the following plausible cross-contextual principles make sense, which constrain the possible variation. Take two individuals x and y in context c such that $M, c \models T(x) \wedge \neg T(y)$. We now constrain the set of contexts C by the following three principles: No Reversal (NR), which forbids x and y to change roles in other contexts:

(NR) $\neg\exists c' \in C : M, c' \models T(y) \wedge \neg T(x)$.

This constraint allows both x and y to be tall in larger contexts than c. However, once we look at such larger contexts, the Upward Difference (UD) constraint demands that there should be at least one difference pair:

(UD) $\forall c' \in C : c \subseteq c' \rightarrow \exists z_1, z_2 : M, c' \models T(z_1) \wedge \neg T(z_2)$.

[1]But see Von Stechow (1984) for an argument saying that also the degree-based approach is in line with Frege's principle.

The final Downward Difference (DD) principle constrains in a very similar way what is allowed if we look at subsets of c: if x and y are elements of this subset, there still should be a difference pair:

(DD) $\forall c' \in C : (c' \subseteq c \;\&\; x, y \in c') \rightarrow \exists z_1, z_2 : M, c' \models T(z_1) \wedge \neg T(z_2)$.

If we say that 'John is P-er than Mary' iff there is a context c s.t. $M, c \models P(j) \wedge \neg P(m)$, Van Benthem shows that the comparative (given the above constraints on context strucures) as defined above has exactly those properties which we intuitively want for most comparatives (see below).

In the definition of a context structure we used above, context structures give rise to orderings for any context dependent adjective. For convenience, we will just limit ourselves to one adjective: P. If we do so, we can think of a context structure as a triple $\langle X, C, P\rangle$, where P can be though of as a *choice function*, rather than as a general valuation function.

Definition 1.
A <u>*choice structure*</u> *M is a triple $\langle X, C, P\rangle$, where X is a non-empty set of individuals, the set of contexts, C, consists of all finite subsets of X, and the choice function P assigns to each context $c \in C$ one of its subsets.*

Notice that $P(c)$ (with respect to context structure M) corresponds to the set $\{x \in X : M, c \models P(x)\}$ in our earlier formulation. To state the cross contextual constraints somewhat more compactly than we did above, we define the notion of a *difference pair*: $\langle x, y\rangle \in D_P(c)$ iff$_{def}$ $x \in P(c)$ and $y \in (c - P(c))$. Now we can define the constraints as follows (where c^2 abbreviates $c \times c$, and $D_P^{-1}(c) =_{def} \{\langle y, x\rangle : \langle x, y\rangle \in D_P(c)\}$):

(NR) $\forall c, c' \in C : D_P(c) \cap D_P^{-1}(c') = \varnothing$.
(UD) $c \subseteq c'$ and $D_P(c) \neq \varnothing$, then $D_P(c') \neq \varnothing$.
(DD) $c \subseteq c'$ and $D_P(c') \cap c^2 \neq \varnothing$, then $D_P(c) \neq \varnothing$.

If we say that $x >_P y$, iff$_{def}$ $x \in P(\{x, y\})$ and $y \notin P(\{x, y\})$, van Benthem (1982) shows that the ordering as defined above gives rise to a strict *weak order*.[2,3] A structure $\langle X, R\rangle$, with R a binary relation on X, is a strict weak

[2]Notice that in contrast to Klein (1980), van Benthem (1982) does not existentially quantify over comparison classes. One can show that given van Benthem's constraints, this doesn't really matter. However, it is possible to slightly change the constraints such that the difference becomes important. In van Rooij (2010a), it is argued that if one does so, one can make a natural distinction between so-called explicit ('John is taller than Mary') and implicit ('John is tall compared to Mary') comparatives.

[3]For the construction of weak orders based on 'maximizing' choice functions, see van

order just in case R is irrreflexive (IR), transitive (TR), and almost connected (AC):

Definition 2.
A (strict) weak order is a structure $\langle X, R\rangle$, with R a binary relation on X that satisfies the following conditions:
(IR) $\forall x : \neg R(x,x)$.
(TR) $\forall x,y,z : (R(x,y) \wedge R(y,z)) \rightarrow R(x,z)$.
(AC) $\forall x,y,z : R(x,y) \rightarrow (R(x,z) \vee R(z,y))$.

If we now define the indifference relation, 'I', or in our case '$\sim_P$', as follows: $x \sim_P y$ iff$_{def}$ neither $x >_P y$ nor $y >_P x$, it is clear that '$\sim_P$' is an equivalence relation. Because of the latter fact, (strict) weak orders play an important role in the theory of measurement.

The idea of measurement theory (Krantz et al, 1971) is that we represent properties of and relations between elements of certain abstract ordering structures in terms of properties of and relations between real numbers that we already understand much better. A *quantitative* measure based on an ordering like $\langle X, R\rangle$ is a representation of the qualitative ordering relation (like *taller than*, represented by 'R') in terms of the quantitative ordering relation *greater than* , '>', between real numbers (**R**). Measurements normally start with (strict) weak orders. It is well-known that such orders R can be turned into *linear* orderings R^* of equivalence classes of individuals that are connected ($\forall v, z : vR^*z \vee zR^*v$). First, say that x is R-equivalent to y, $x \sim_R y$, iff$_{def}$ neither xRy nor yRx. Take $[x]_{\sim R}$ to be the equivalence class $\{y \in X | y \sim_R x\}$. Then we say that $[x]_{\sim R} R^* [y]_{\sim R}$ iff$_{def}$ xRy. Measures – i.e., real numbers – are then assigned to elements of these equivalence classes. A measurement is defined in terms of a (homomorphic) function f that assigns each element of X to a real number such that $\forall x, y \in X : xRy$ if and only if $f(x) > f(y)$ and to each $y \in [x]_{\sim R} : f(y) = f(x)$. In general, there are many alternative mappings to f that would numerically represent the qualitative relation R equally well. However, this mapping f is unique *up to a certain group of transformations*. For instance, the mappings f and g to the (ordered) set of real numbers represent the same (*ordinal*) ordering structure $\langle X, R\rangle$ in case they can be related by a (strictly) *monotone increasing transformation*: for any $x, y \in X : f(x) > f(y)$ iff $g(x) > g(y)$. Quantitative scales that are unique up to such strictly monotone increasing transformations are called **ordinal scales**. Notice that in scales that are just ordinal scales, the numbers represent nothing more than the 'greater than' relation. Thus, ordinal measurements represent nothing more than qualitative orderings. In particular,

---

Rooij (2010b).

the resulting numerical ordering is *not* a degree structure as invoked by traditional degree-based analyses of comparatives. A prominent example of an ordinal scale used by linguists is Lewis's (1973) comparative similarity ordering used in the analysis of counterfactual conditionals.

3 Degrees: interval and ratio-scales

It is generally agreed (e.g. von Stechow, 1984) that there are at least two types of examples for which more is required than just qualitative orderings and we need degrees: (i) examples like (2-a), (2-b) and (2-c) that explicitly talk about degrees, and (ii) examples like (2-d) (and perhaps (2-e)) that involve two different types of adjectives:

(2)
- a. John is 5 centimeters taller than Mary.
- b. John is twice as tall as Mary.
- c. John is 1.80 meters tall.
- d. John is taller than Mary is wide.
- e. John is more happy than Mary is sad.

To account for degrees we need measurement theory. We have seen above that a quantitative measure based on a linear order $\langle X, R\rangle$ is defined in terms of a (homomorphic) function f that assigns each element of X to a real number such that $\forall x, y \in X : xRy$ if and only if $f(y) > f(z)$. If we just want to represent an ordinal scale, the set of mappings that would do this job would just have to be unique op to strictly monotone increasing transformations. To faithfully represent *more informative* ordering structures, however, the different mappings should be unique up to a *larger* group of transformations.[4]

Suppose, for instance , that we have a collection of straight sticks of various sizes and we assign a number to each stick by measuring its length using a ruler. If the number assigned to one stick is greater than the number assigned to another stick, we can conclude that the first stick is longer than the second. Thus a relationship among the numbers (greater than) corresponds to a relationship among the sticks (longer than). This is all that is required for ordinal scales. But measurement of length satisfies another constraint as well. Suppose that we lay two sticks end-to-end in a straight line and measure their combined length, then the number we assign to the concatenated sticks will equal the sum of the numbers assigned to the individual sticks. Now an additional relationship among the numbers (addition) corresponds to an additional relationship among the sticks, that of concatenation. Thus, to faithfully represent this additional relationship among the sticks, the measures have to

[4]Where the notion of 'group' is used in the algebraic way.

be related in an additional way as well. This additional constraint is captured by demanding uniqueness with respect to a larger group of transformations.[5]

Ratio scales Suppose that the ordering structure can make sense not only of sentences like 'x is taller than y', but also of sentences like (2-b) 'John is twice as tall as Mary.' and (2-c) 'John is 1.80 meter tall'. In this case we need a theory of degrees that involves *addition* or *multiplication*. Standard measurement theory gives us that for so-called *extensional* or *additive magnitudes*, and the corresponding scales are called *ratio scales*. Standard examples of extensional magnitudes are *length*, and *duration*. The corresponding algebraic systems are called *closed extensive structures*. (Qualitative) ordering structures that give rise to extensional measurement must come with an operation of *concatenation*, a specified procedure for joining two objects, denoted by '$\circ$'. Thus, the ordering structures must be of the form $\langle X, R, \circ\rangle$, where operation '$\circ$' satisfies certain conditions, like associativity.[6]

For mapping f to faithfully represent such an ordering structure it must be the case that for all $x, y \in X : f(x \circ y) = f(x) + f(y)$ (of course, it must also hold that xRy iff $f(x) > f(y)$). Just as before, this mapping is unique up to a certain group of transformations. The group of transformations are now all of the form $g(x) = \alpha f(x)$, with $\alpha > 0$, which are known as *similarity transformations*. The existence of α means that the unit of measurement is arbitrary, but in extensional magnitudes we measure with respect to an absolute zero-point. It is worth to point out that the scale structures invoked in standard degree-based approaches (e.g. Kennedy, 1999, von Stechow, 1984) are generally ratio-scales.

Interval scales Ratio-scales are very rich, but also presuppose a lot, in particular, an absolute zero-point. But such absolute zero-points cannot always be assumed. Still, even though we cannot assume an absolute zero-point, we still want to interpret certain types of sentences for which we require more than just an ordinal scale. For instance, suppose that the ordering structure

[5]In the following I will discuss only ordinal, ratio and interval scales, the scales considered by Stevens (1946). Krantz et al. (1971) distinguish other scales as well, such as Log-interval, difference, and absolute scales. Although also these scales are important to represent psychophysical functions (Log-interval and difference scales) and probability (absolute scales), I will ignore them in this paper.

[6]The triple $\langle X, \geq, \circ\rangle$ is a *closed extensive structure* iff (i) $\langle X, \geq\rangle$ is a weak order, i,e., '$\geq$' is a reflexive, transitive, and connected relation; (ii) '$\circ$' satisfies weak associativity: $x \circ (y \circ z) \sim (x \circ y) \circ z$; (iii) '$\geq$' and '$\circ$' satisfy Monotonicity: $x \geq y$ iff $x \circ z \geq y \circ z$ iff $z \circ x \geq z \circ y$, and (iv) has the Archimedean property: If $x > y$, then for any $v, w \in X$, there exists a positive integer n such that $nx \circ v \geq ny \circ w$, where nx is defined inductively as $1x = x, (n_1)x = nx \circ x$.

can make sense not only of sentences like 'x is taller than y', but also of things like (2-a) and (3):

(3) x is P-er than y *by more than* v is P-er than w.

In measurement theory this is standardly done by taking a quaternary relation D on X to be basic, and define the corresponding binary relation in terms of it. Ordering structures in which such quaternary relations are taken as basic are called 'algebraic difference structures'. For a mapping f to represent this latter type of information, it should be the case that $xyDvw$ iff $f(x) - f(y) > f(v) - f(w)$. Algebra shows that $a - b \geq c - d$ iff $(\alpha a + \beta) - (\alpha b + \beta) \geq (\alpha c + \beta) - (\alpha d + \beta)$ (where $\alpha > 0$). This means that the ordering of *intervals* between two numbers is preserved under a positive linear transformation (meaning that (2-a) becomes meaningful). In fact, the *ratios* between the intervals are preserved as well.[7] A natural proposal (Krantz et al, 1971) is then to assume that two mappings f and g faithfully represent the same such an ordering structure, if f and g are the same up to a strictly *positive linear transformation*, i.e., for any $x \in X : g(x) = \alpha(f(x)) + \beta$, where α, β are real numbers, and $\alpha > 0$. The number α represents the fact that the *unit* of measurement is arbitrary, like in ratio-scales. The difference between interval-scales and ratio-scales is the use of the additional β. The number β represents the fact that the ordering structure doesn't have a fixed zero-point. But such fixed zero-points are required to make sense of examples (2-b) and (2-c). Quantitative scales that are unique up to such strictly positive linear transformations are called *interval scales*, and the magnitudes measured this way *intensional magnitudes*. Thus, in terms of interval scales we can interpret examples like 'x is taller than y', (2-a), and (3), but not examples like (2-b) and (2-c). Three well-known magnitudes that gives rise to an interval scale (i.e., for which we need an interval-scale but don't want a ratio-scale) are 'utility', '(clock)time' and 'warmth' (forgetting about Kelvin, for a minute). We can measure temperature, for instance, in terms of degrees Celsius and degrees Fahrenheit, and it is well-known that we can transform the one to the other by means of a positive linear transformation: x degrees Celsius is $\frac{9}{5}x + 32$ degrees

[7]Suppose that g is obtained by a positive linear transformation from f. This means that $g(x) - g(y) = \alpha x + \beta - \alpha y - \beta = \alpha(x - y)$, for some $\alpha > 0$. Similarly $g(v) - g(w) = \alpha(v - w)$. But this means that $\frac{g(x)-g(y)}{g(v)-g(w)} = \frac{f(x)-f(y)}{f(v)-f(w)}$. Thus, we can *also* make sense of comparatives like the following: 'x is P-er than y by *2 times* more than v is P-er than w'. At first it seems possible to make sense of differences between intervals such that (3) is meaningful, but still not make the above comparative meaningful. Scott and Suppes (1958) show, however, that this is not the case if we want to reduce the direct measure of intervals like xy to the direct measures of the individuals x and y themselves ($f(xy) = f(x) - f(y)$).

Fahrenheit, and x degrees Fahrenheit is $\frac{5}{g}x + (-\frac{32\times5}{9})$ degrees Celcius.[8] Notice that one degree Celsius warmer is not the same as one degree Fahrenheit warmer (but as $\frac{9}{5}$ degrees Fahrenheit warmer), and that 0 ^oC is not at all the same as 0 ^oF, but rather as 32 ^oF.

It is clear that the distinction between ordinal scales and ratio-scales has linguistic relevance. But this also holds for the interval vs. ratio distinction, because there arguably are some linguistic examples that require an interval-scale to make them meaningful, but for which we don't want a ratio-scale. Two such examples are it negative adjectives and *similarity* (or difference) *comparisons*. As for negative adjectives, suppose we want to give a representation in terms of an order preserving real valued function f not only of the ordering 'taller than', $\langle X, >_T\rangle$, but at the same time also of the relation 'shorter than'. How should this function look like? There has been a lot of discussion of this. According to Creswell (1976) and Klein (1980), one's shortness is the same as one's tallness, it is just that the order is reversed. According to Bierwisch (1984), one picks out the average, or prototypical value, p, and measures the shortness of x as $p - f_T(x)$, where $f_T(x)$ gives x's tallness. According to a very prominent analysis (going back to Seuren (1984) and von Stechow (1984), but most forcefully defended by Kennedy, 1999) the shortness of x is $\infty - f_T(x)$, where ∞ is infinity. All of these approaches assume that 'shortness' should be measured in terms of a ratio-scale, and none of them is uncontroversial. Recently, Sassoon (2010) observed that to account for the data, Kennedy does not need to make use of the unreal number ∞ to substract $f_T(x)$ from: any real number n would do. Also if the shortness of x is not $\infty - f_T(x)$ but rather $n - f_T(x)$, one can explain why 'y is 2 cm shorter than x' is true iff 'x is 2 cm taller than y' is true. Moreover, she points out that if the choice of the real number n (i.e., the zero point) is left undetermined, one can explain why the sentences 'y is 3 times as short as x' and 'x is 180 cm short' are not meaningful. Formally, the reason is that the resulting 'shorter than' relation gives rise to an interval scale, which does not have an absolute minimum: if $f_S(y) = -6 = -f_T(y)$, for instance, and $f_S(x) = -2 = -f_T(x)$, then $f_S(y)$ is three times as much as $f_S(x)$. But from this it doesn't follow that we can thus meaningfully conclude that y is three times as short as x, if shortness gives (only) rise to an interval scale. To warrant this conclusion it has to be the case that this ratio is preserved under any positive linear transformation. But this is not the case: take for any z, $g_S(z) = f_S(z) + 1$. It follows that $g_S(y) = -5$ and $g_S(x) = -1$, meaning that $g_S(y)$ is as much as 5 times as much as $g_S(x)$. The mapping $g_S(\cdot)$ can obviously be defined as proposed by Sassoon (2010): $g_S(z) = -f_T(z) + 1$. Although it still holds that 'x is 4 cm shorter than y' is

[8] Notice that the mapping from degrees Fahrenheit to degrees Celcius is the inverse of the mapping from Celcius to Fahrenheit.

true iff 'y is 4 cm taller than x' is true, the claim that 'y is 3 times as short as x' is not true anymore. Thus, ratio-claims are correctly predicted not to be meaningful if 'shortness' is measured in terms of an interval scale.

Consider the following similarity statement:

(4) John is more similar (different) to (from) Mary than Bill is to (from) Sue.

In psychology, the notion of similarity plays an important role to account for categorization. The two dominant models in this field are (i) the so-called geometrical model (e.g. Shephard, 1962), and (ii) Tversky's (1977) contrast model. In the geometrical model, individuals are mapped into some (Euclidean) space, and their similarity is measured in terms their closeness in this space. Although this model is very natural for the representation of some features, such as colors, it is much less natural to represent the similarities of, for instance, natural kinds. Tversky's (1977) contrast model seems more suitable here. Tversky (1977) measures the similarity of two individuals in terms of the sets of (qualitative) features they share and which they differ on. Tversky proves that his notion of similarity gives in general rise to an interval scale. And that seems appropriate: statements like 'x is similar to y to degree n' and 'x is *twice as similar* to y than v is to w' only make sense in exceptional circumstances.

4 Interadjective comparisons

We want to make sense of some interadjective comparatives like

(1) x is P-er than y is Q.

As before, we need particular structures to do this. But we don't necessarily want these structures to also support comparatives involving magnitudes of difference or examples like (5).

(5) x is *three times* P-er than y is Q.

One prominent recent account (Bale, 2008) predict these should be ok;[9] two alternative analyses, based on extensions of the standard measurement and the delineation-based approach developed earlier in this paper, support a more limited set of comparative statements. I will argue that these alternative analyses

[9] Just like the fuzzy logic account of comparatives.

are much more general, and sometimes more accurately capture the linguistic facts.

4.1 Bale's analysis of indirect comparison

In linguistics it is generally assumed that to make comparatives like (1), we *have* to make use of degrees. Either because P and Q are commensurable adjectives like 'long' and 'wide' (their orders have the same zero-point and the same unit of measurement), or the comparative is seen as an *indirect* or *relative* comparison (Bale, 2008) where the rank of the two items is compared in their respective 'P-er than' and 'Q-er than' orderings. Thus, if we limit ourselves to x, y and z, and $y >_P x >_P z$ and $z >_Q x >_Q y$, then (1) is true on its indirect reading.[10,11] Bale argues that relative comparison is also relevant for cases like the following:

(6) Ella is heavier for a baby than Denis is for a three year old.

And indeed, also here direct comparison doesn't make a lot of sense.

Bale's (2008) account of relative comparisons starts out with a set of (linear) orderings $\langle X, >_P\rangle, \langle Y, >_Q\rangle$, etc. From this he constructs a universal ordering out of this as follows. First, he defines the *bale-rank* of $x \in X$ with respect to the ordering $\langle X, >_P\rangle$ as the amount of (equivalence classes of) individuals ordered below x: $b_P(x) =_{df} |\{[y]_P \in X : x >_P y\}|$, where $[y]_P$ is an equivalence class with respect to $>_P$ ($z \in [y]$ iff $y \not>_P z$ and $z \not>_P y$). Then he proposes to order individual-property pairs in terms of their bale-ranks: $\langle x, P\rangle > \langle y, Q\rangle$ iff $b_P(x) > b_Q(y)$.[12] Of course, this doesn't make much sense, unless the orderings $\langle X, >_P\rangle, \langle Y, >_Q\rangle$, etc. give rise to equally many ranks. This indicates that Bale's proposal is very limited. However, there exists a trick to guarantee that the orderings have equally many ranks: we can make use of *normalization*: the bale-rank of x given ordering $\langle X, >_P\rangle$ depends on the number of ranks of this ordering: $b_P(x) =_{df} \frac{|\{[y]_P : x >_P y \ \& \ y \in X\}|}{|[x]_P : x \in X|}$. But for this normalization to make any sense (to be well-defined), it is required that the number of equivalence classes is finite, meaning that the original order has to have a minimal and a maximal

[10]Kennedy (1999) claims that only 'direct' comparisons make sense, or ones reducible to such direct ones. Though I agree that not everything is possible, I agree with Bale and many others that Kennedy's constraint is too strict.

[11]Bale (2008) argues that from the truth of the comparative 'x is P-er than y is Q' it does not follows that x is P and that y is Q. It certainly doesn't follow from 'the table is higher than it is wide', but also not from 'That baby is taller than it is old', or so I would say.

[12]The bale-rank is really the same as the Borda-count in social choice theory. One reviewer suggested that Bale's construction is a special case of the Mostovski collaps in set theory. If that is true (I am not sure), linearity of the orders is not essential, but well-foundedness is.

element.[13] Thus, our normalization trick can only be used in a very limited set of circumstances.

What kind of scale does Bale (2008) end up with? It is easy to see that he ends up with a *ratio scale*.[14] Indeed, the resulting ordering gives rise to an obvious 0 and 1: $b_P(x) = 0 = b_Q(y)$ in case x is lowest in ordering $\langle X, >_P \rangle$ and y lowest in ordering is lowest in $\langle Y, >_Q \rangle$, and $b_P(x) = 1$ in case x is the highest in the former ordering. Recall that if the ordering $\langle X, >_P \rangle$ gives rise to a ratio-scale it makes sense to make comparisons like 'x is *three times* as P as y'. This, of course, is thus immediately predicted by Bale: statements like 'x is *three times as* P as y is Q' make perfect sense. Apart from the fact that it is not obvious that such interadjective comparisons always make sense (see section 4.2), it is also surprising, given that he started out (possibly) with two ordinal scales $\langle X, >_P \rangle$ and $\langle Y, >_Q \rangle$. A similar objection goes as follows: It it is only natural to assume that a standard comparative like 'x is P-er than y' is just a special case of the subdeletion comparative-construction 'x is P-er than y is Q'. It is special, because with standard comparatives the two adjectives involved are the same. From this it follows that if restricted to adjective P, the two orderings should be the same. And indeed, it follows by Bale's construction that 'x is P-er than y' is true with respect to ordinal scale $\langle X, >_P \rangle$ if and only if it is true on the universal ordering. However, if the 'direct' P-measure gives rise to a ratio-scale, the relative P-measure of any object x is in general not the result of a linear transformation of its direct measure.[15] Moreover, if the 'direct' measure only gives rise to an ordinal scale, the universal ordering allows for much more specific comparatives. Although the $\langle X, >_P \rangle$ was only an ordinal scale, the universal ordering was a ratio-scale, and thus contains enough information to evaluate comparatives like 'x is *three times as* P as y'. Thus, this comparative is meaningful on the universal scale, though not on its original scale: a surprising consequence. Indeed, if we take into account two adjectives of which (at least) one only gives rise to an ordinal scale, the combined scale cannot be more than an ordinal scale.

These last criticisms, however, are not as serious as it might seem. The reason is that Bale (2008) actually started out with more than suggested: he started out with two orderings that determine the *relative* P- and Q-ness. These orderings, however, give rise to *ratio*-scales. Bale assumes that 'x is higher than y is wide' can be interpreted in (at least) two different ways. One way to account for this comparative is in terms of what Bale calls 'direct' measurement (we measure the height and width in terms of centimeters, and then compare). On a second reading the evaluation of this comparative de-

[13]I owe this observation to an anonymous reviewer.

[14]At least, if $\langle X, >_P \rangle$ and $\langle Y, >_Q \rangle$ are known.

[15]I owe this observation to an anonymous reviewer.

pends more on the *relative* length and width. But to make sense of this claim, and not to fall in the above described problem, Bale also has to assume that standard comparatives like 'x is three times as high as y' can be given such a relative reading. Similarly for adjectives that normally are assumed only to give rise to an ordinal scale. The question is whether this consequence is desired. The major difficulty of Bale's approach, however, is clearly that it can be used only in a very limited set of circumstances: the different orderings musst have equally many ranks. One wonders whether it is not possible to account for interadjective comparisons in other circumstances as well.

4.2 A measurement-theoretic analysis of interadjective comparisons

If we concentrate on comparisons that involve different types of adjectives, we can see that there is a crucial difference between what is required to make sense of (7) and (8):

(7) John is taller than Mary is smart.

(8) John is 5 cm taller than Mary is wide.

It is obvious that the purely qualitative analysis of Klein (1980) discussed above cannot account for (8). But even if we make use of explicit measures, the requirement on how specific one should think of degrees to make make sense of (8) is quite different from the 'type' of degree one needs to account for (7). Suppose we can make sense of (7). Does it follow from this that we can thus also make sense of (9)?

(9) ?John is *5 times* as tall as Mary is smart.

We have seen that according to Bale (2008) the answer must be positive. The question is whether it needs to be.

Recall that according to Bale an analysis of indirect comparison should also be able to account for (6).

(6) Ella is heavier *for a baby* than Denis is *for a three year old.*

It is, I believe, an interesting observation to see that the above type of comparison is virtually identical to the one in (10) which plays a major role in political economy:

(10) Action x is more useful *for John* than action y is *for Mary.*

More in general, the issue of interadjective comparison is structurally analogous to the issue how to account for intersubjective comparison of utility in political economy and social choice theory (see Sen, 1970 for overview).[16] The different subjects (e.g. utilities for John and for Mary) in economics play the role analogous to the different adjectives in cases we consider. This observation is important, because by analogy we can learn from political economy what kinds of measurement-theoretic structures we need to account for particular types of interadjective comparisons.

Although there exists a large group of economists who claim that statements like (10) make sense, almost nobody claims that (11) is sensible:

(11) Action x is *5 times* as useful for John than action y is for Mary.

On the other hand, the meaningfulness of the following type of statement is disputed by far less political economists:

(12) Action x is more useful for John than y *by more* than action v is more useful for Mary than action w.

Notice that the required conditions for a statement like (12) to be meaningful are similar to the conditions under which an interadjective comparison statement like (13) to be meaningful:

(13) John is taller than Mary *by more* than this room is warmer than that room.

To make sense of the difference in meaningfulness between these types of comparisons, we can make use of measurement theory again.

We have seen above that measurement theorists distinguish *ordinal* versus *interval* versus *ratio* scales in terms of constraints of transformations. The scales are based on the assignment of a measure to a set of individuals. The crucial idea to account for interadjective comparison is to define scales based on the assignment of measures to a set of *individual-adjective pairs.* In the previous sections we assumed that to represent an ordering $\langle X, > \rangle$ numerically, we define an order-preserving function f from X to the set of real numbers $\mathbf{R}$. To make sense of interadjective comparisons, we can think of the domain not just as a set of individuals, but rather as a set of individual-adjective pairs. Thus, if Ad is the set of relevant adjectives, we should now think of f as a function from $X \times Ad$ to $\mathbf{R}$. Once we have such a function, we can make sense of interadjective comparisons. But, of course, assuming any such function is

[16] The issue of how to make sense of interpersonal comparisons of utility is taken to be important in social choice theory, because it allows in a natural way to save the theory from Arrow's (1950) impossibility results (cf. Harsanyi, 1955; Sen, 1970).

rather ad hoc, and as before, we want this function to be invariant under some transformations. But not any transformation will do: if $f(\langle x,P\rangle) = 2$ and $f(\langle y,Q\rangle) = 1$, the transformation that maps any $f(\langle z,S\rangle)$ into $-f(\langle z,S)$ will reverse the ordering, which is something that we don't want.

Let us first look at the transformations that preserve the *ordinal* structure of interadjective comparisons: the type of transformations that already allow for comparisons like (2-d) and (7). The only thing that is needed is that the function f from $X \times Ad$ to $\mathbf{R}$ should be unique up to strictly monotone increasing transformations. Thus, it should be the case that $f(\langle x,P\rangle) > f(\langle y,Q\rangle)$ iff $g(\langle x,P\rangle) > g(\langle y,Q\rangle)$, i.e., if g behaves monotone increasingly w.r.t. f. Thus, we can think of g as defined in terms of a strictly monotone increasing function φ from real numbers to real numbers such that $g(\langle x,P\rangle) = \varphi(f(\langle x,P\rangle))$ for all $x \in X$ and $P \in Ad$. We might call this type of ordinal interadjective ordering relation a *co-ordinal scale*.

The types of transformation that preserves differences, and ratios between differences, and thus allow for comparisons like (13) (and (12)) are the strictly positive linear transformations. Thus, the mapping g from $X \times Ad$ to $\mathbf{R}$ gives rise to the same differences, and ratios between differences, as mapping f iff g can be derived from f by means of a positive linear transformation: $\forall\langle x,P\rangle \in X \times Ad : g(\langle x,P\rangle) = \alpha \times (f(\langle x,P\rangle)) + \beta$, where α, β are real numbers, and $\alpha > 0$. Notice that neither β nor α depends on the adjective. Let us call this type of interadjective ordering relation a *co-interval scale*.[17] Observe that just as interval-scales are all that is needed to account for examples like (2-a), co-interval scales are all that is needed to make comparatives like (8) meaningful.[18]

The types of transformation, finally, for which comparisons like (9) make sense as well are those that also preserve addition and multiplication. The idea of such a transformation is that it can be done with respect to an *independent norm*. Typically, this norm is the *minimum*, or zero. Thus, now the mapping g from $X \times Ad$ to $\mathbf{R}$ is considered to be 'the same' as mapping f iff g can be derived from f by means of a similarity transformation: $\forall\langle x,P\rangle \in X \times Ad$: $g(\langle x,P\rangle) = \alpha \times f(\langle x,P\rangle)$, for some $\alpha > 0$. In analogy with what we did before, we will call this type of ordering a *co-ratio* scale. Notice that length and width can be compared in this strong sense, because they share their minimum, but that this is not the case for length and intelligence, for instance.[19]

[17] In fact, there is a different constraint as well that still might be called 'co-interval': g is 'the same' as mapping f iff $\forall\langle x,P\rangle \in X \times Ad : g(\langle x,P\rangle) = \alpha \times (f(\langle x,P\rangle)) + \beta_P$, where α, β_P are real numbers. This time β_P *is* dependent on the adjective P.

[18] Which doesn't mean that we ever make such comparisons if we can only construct a co-interval scale.

[19] At least, if not taking in their *relative* sense as meant by Bale.

In general, the set of measurement functions that take as input all individual-adjective pairs and count as 'the same' will probably not be closed under many transformations.[20] However, if we limit the functions to proper subsets of $X \times Ad$ it might, and which transformations it satisfies will depend on the limitation we consider. It seems natural to assume that if we limit ourselves to the adjectives 'tall' and 'wide' measured in an absolute way, for instance, we end up with a ratio-scale. Thus, if we limit ourselves to the dimensions height, width, and depth, we can assume that the set of measurement functions F is closed under similarity transformations. If we take other adjectives, or dimensions, into account as well for which interadjective comparisons still make sense, we can assume a different set of measurement functions H closed under strictly monotone increasing transformations (or even positive linear transformations). If not all interadjective comparisons make sense, this set H does not include the whole set $X \times Ad$ as its domain.

The set of measurement functions closed under similarity transformations might be larger than one might think initially. Making use of normalization, I think that the following types of comparison should give rise to a ratio-scale:

(14) a. This table is more red than that that chair is blue.
b. A bat is more of a bird than a whale is a fish.

What seems to go on in (14-a), for example, is that we measure the redness of this table, and the blueness of that chair, in terms of the similarity (after normalization) with the prototypes of red and blue. One can make use of a ratio-scale in this case because the prototypes function as the obvious minima (maximal similarity is minimal difference). Maxima can be determined similarly. This is very close to the comparison of deviation cases discussed by Kennedy (1999). Bale shows that we can end up with co-ratio-scales as well if we start out with the *relative* positions in the respective scales. But we have seen that at least a parallel case of intersubjective utility comparison is taken to be at most co-interval according to economists. And given the scale-structures of tallness and wisdom (they don't have maximal elements), Bale cannot account for a parallel example like 'x is taller than y is wise'. It doesn't seem to be a major claim to say that many interadjective comparisons are at most co-ordinal. Out of context, numbers assigned to x's tallness and y's wisdom to be compared with each other don't seem to represent much.

We have seen how to *characterize* co-ordinal scales, but how can we *construct* one? We have seen that it is easy to construct *rich* interadjective measurements. But how can one construct a notion of interadjective measurement

[20]Though this might depend on whether we want the comparative to be 'direct' or 'indirect' in the sense of Bale (2008).

that gives rise to weaker scales only?

4.3 Constructing co-ordinal and co-interval scales

Bale started out with a number of ordinal scales, and wondered how to turn this into a universal scale. Bale concluded that this is possible only if we assign *numbers* to the ranks. However, it seemed that if we want to make sense of comparatives involving two adjectives that themselves only give rise to ordinal scales we do *not want* numbers. In the previous section we have seen that it is possible to do so by making use of co-ordinal measurements. But how can such a measurement be constructed? In this section I will sketch two possible ways to go. The first construction will be very similar to what we did in section 2, but this time involves not just individuals, but rather individual-property pairs. The second construction is really due to Klein (1980), following McConnell-Ginet (1973).

In section 2 we constructed a 'P-er than' ordering out of a choice structure. Now we will suggest that exactly the same idea can be used to construct a interadjective ordering, if we take the objects to be individual-adjective pairs, instead of individuals. Let $X^* = X \times Ad$, where X is a set of individuals and Ad a set of adjectives or properties. Assume that C is a set of finite subsets of X^*. We make use of the abstract 'choice function', 'Lots', and say that $\langle x, P\rangle \in Lots(c)$ means that x has lots of P-ness compared to the properties other individuals have in c. We assume that the constraints on how 'choice function' '*Lots*' behaves among different comparison classes (of individual-adjective pairs) are just the same as the constraints that choice function P had to obey on comparison classes of individuals in section 2. If we then say that $\langle x, P\rangle > \langle y, Q\rangle$ iff $\langle x, P\rangle$ is an element of $Lots(\{\langle x, P\rangle, \langle y, Q\rangle\})$, but $\langle y, Q\rangle$ is not, it immediately follows that the ordering relation between individual-adjective pairs will be a strict weak ordering, giving rise just to an *ordinal* scale. In terms of measurement theory, it still holds that g would be an equally good representation of the order as f, $g \approx f$, iff $\forall \alpha, \beta \in I \times Ad : f(\alpha) > f(\beta)$ iff $g(\alpha) > g(\beta)$. But if we want to make sure that it doesn't make much sense to say things like 'x is taller than y is intelligent', we cannot assume that all the intercontextual constraints of section 2 are obeyed in general. Perhaps in general we want the ordering relation only to be a strict partial order,[21] and perhaps even less (e.g. only acyclic).[22,23]

[21] A relation R gives rise to a strict partial order iff R is (i) irreflexive ($\forall x : \neg xRx$), and (ii) transitive ($\forall x, y, z : xRy \wedge yRz \rightarrow xRz$).

[22] A relation R is acyclic iff there are no $x_1, ..., x_n$ such that $x_1Rx_2, x_2Rx_3, ..., x_{n-1}Rx_n$ and x_nRx_1. Notice that R is irreflexive and asymmetric if it is acyclic.

[23] For a discussion of how to construct (among others) strict partial orders and semi-orders

In the above construction I assumed that whether or not $\langle x, P\rangle \in \mathit{Lots}(c)$ holds is a *primitive* fact of our model. But perhaps one would like to determine its truth value in a somewhat different way.

Although Klein (1980) is best known for his comparatives as sketched in section 2, he already suggested an analysis of subdeletion comparatives that is somewhat richer.[24] Rather than quantifying over comparison classes, he existentially quantifies over (the meanings of) *modifiers of adjectives*, like *very* and *fairly*. One motivation for quantifying over such modifiers is to be able to account for subdeletion comparatives like (15-a), which are interpreted as something like (15-b) as suggested earlier by McConnell-Ginet (1973).

(15) a. John is more happy than Mary is sad.
b. $\exists \mathbf{f} \in \{\text{very, fairly, quite, ...}\}[\mathbf{f}(\mathit{Happy})(j) \wedge \neg \mathbf{f}(\mathit{Sad})(m)]$.

Klein (1980) accounts for modifiers of adjectives in terms of comparison classes and shows that existentially quantifying over comparison classes is only a special case of quantifying over these modifiers. To illustrate this, suppose we have a set of 4 individuals: $X = \{w, x, y, z\}$. One comparison-class, call it c_0, is X. Suppose now that $P(c_0) = \{w, x\}$, and (thus) $\overline{P(c_0)} = \{y, z\}$ (with $\overline{Y}$ as the complement of Y with respect to the relevant comparison class). We can now think of $P(c_0)$ and $\overline{P(c_0)}$ as new comparison classes, i.e., $P(c_0) = c_1$, and $\overline{P(c_0)} = c_2$. Let us now assume that $P(c_1) = \{w\}$ and $P(c_2) = \{y\}$. If so, this generates the following ordering via Klein's definition of the comparative we used before: $w >_P x >_P y >_P z$. Let us now assume that $\mathbf{f}$ is a modifier of adjectives. According to Klein's (1980) final analysis, he represents comparatives of the form 'x is P-er than y' as follows:

(16) $\exists \mathbf{f}[\mathbf{f}(P)(c_0)(x) \wedge \neg \mathbf{f}(P)(c_0)(y))]$

To continue our illustration, we can define the following set of modifier functions on domain X in terms of the behavior of P with respect to different comparison classes: $\mathbf{f_1}(P)(c_o) = P(c_0)$, $\mathbf{f_2}(P)(c_o) = P(P(c_0))$, $\mathbf{f_3}(P)(c_o) = P(P(c_0)) \cup P(\overline{P(c_0)})$, and $\mathbf{f_4}(P)(c_o) = c_0$. Thus, $\mathbf{f_1}(P)(c_o) = \{w, x\}$, $\mathbf{f_2}(P)(c_o) = \{w\}$, $\mathbf{f_3}(P)(c_o) = \{w, x, y\}$,[25] and $\mathbf{f_4}(P)(c_o) = \{w, x, y, z\}$. Of course, one can similarly stipulate that $\mathbf{f_1}(Q)(c_o) = Q(c_0)$, $\mathbf{f_2}(Q)(c_o) = Q(Q(c_0))$, $\mathbf{f_3}(Q)(c_o) = Q(Q(c_0)) \cup Q(\overline{Q(c_0)})$, and $\mathbf{f_4}(Q)(c_o) = c_0$ for adjective (choice function) Q. Take $\mathbf{F}$ to be $\{\mathbf{f_1}, \mathbf{f_2}, \mathbf{f_3}, \mathbf{f_4}\}$.[26] If we limit ourselves to adjective P, this new

based on a Klein-based analysis of comparatives, see van Rooij (2009).

[24]This paragraph comes from van Rooij (2008). More recently, Doetjes (2009) argued in favor of this construction of Klein (1980) to account for comparisons in general.

[25]Thanks to a reviewer for spotting an earlier mistake.

[26]Observe that if we allow the starting comparison classes to vary, it is straightforward to

analysis of the comparative gives rise to the same order: $w >_P x >_P y >_P z$. Moreover, any of those comparatives can only be true according to the new analysis, if it is true according to the old analysis: The statement '$w >_P x$' is true, for instance, because of function $\mathbf{f_2}$. But $\mathbf{f_2}(P)(c_0)(w) \wedge \neg \mathbf{f_2}(P)(c_0)(x)$ holds iff $P(c_1)(w) \wedge \neg(P)(c_1)(x)$, which demonstrates (for this special case) that the old analysis is indeed a special case of the new analysis. What is interesting from our perspective is that Klein's analysis is immediately an analysis of interadjective comparisons based on the meaning of adjectives in different comparison classes. One might argue, however, that the analysis is too general: it seems to predict that all interadjective comparisons make sense, which is a false prediction. This counterargument is based on the assumption that these functions $\mathbf{f}$ always have to be *total*, applicable to *any* adjective. But, of course, one does not need to accept this, if one adopts Klein's proposal: the functions can just be partial, and the domain on any of these functions are just those adjectives for which interadjective comparisons make sense.

It seems to be the case that for many interadjective comparisons, comparatives involving such phrases as '*much* more', '*a bit* more', etc. make sense. This suggests that what is involved is at least a co-interval scale, though one with inexact numbers. Can we think of a construction that gives rise to such a co-interval scale that is not a co-ratio-scale? Klein's construction gives rise, in general, only to an ordinal scale. However, in terms of his construction one could construct something like an interval scale as well. Notice first that according to the above construction, one can *order* the modifiers in terms of what these modifiers do: $\mathbf{f_2} \geq \mathbf{f_1} \geq \mathbf{f_3} \geq \mathbf{f_4}$, because for all $x \in X : \mathbf{f_2}(P)(c_0)(x) \models \mathbf{f_1}(P)(c_0)(x) \models \mathbf{f_3}(P)(c_0)(x) \models \mathbf{f_4}(P)(c_0)(x)$.[27] I suggest that in terms of this extra information one can also account for sentences as (3), (13), and (17).

(17) x is P-er than y *by more than* v is Q-er than w.

Here is the idea: first one can select the set of $\mathbf{f}$s which make 'x is P-er than y' true: $\{\mathbf{f} \in \mathbf{F} : \mathbf{f}(P)(c_0)(x) \wedge \neg \mathbf{f}(P)(c_0)(y))\}$. Call this set $\mathbf{F}(P)(\langle x, y\rangle)$. In a similar way we can define $\mathbf{F}(Q)(\langle v, w\rangle)$. Now one can say that (17) is true iff $\mathbf{F}(Q)(\langle x, y\rangle) \subset \mathbf{F}(P)(\langle x, y\rangle)$. Similarly, one could say that (18),

(18) x is *5 times* as P as y is Q.

is true, for instance, iff the set of $\mathbf{f}$s which make $\mathbf{f}(P)(c_0)$ true[28] is five times

account for comparisons like (6).

[27]It is sometimes claimed that from 'x is P-er than y is Q' it follows that x is P. This could be accounted for by stipulating that for interadjective comparisons we look only at the 'positive' modifiers (in our case, $\mathbf{f_1}$ and $\mathbf{f_2}$).

[28]The use of similar sets has been proposed by Klein (1991) already, to account for com-

as large as the set of **f**s which make $\mathbf{f}(Q)(c_0)$ true. Thus, one could think of the construction as even resulting in a (co)-ratio-scale. This is not meant to be a very serious suggestion, but it just shows that with a little imagination one can construct scales that can – at least in principle – account for (more involved) interadjective comparisons.

5 Multidimensional adjectives

'Clever' and 'big' are usually taken to be multidimensional adjectives. It is then standardly assumed that each 'dimension' gives rise to a separate ordering, and that whether somebody is more clever/big than somebody else depends then, somehow, on these separate orderings. This suggests that our above analysis of inter-adjective comparisons is relevant as well for the analysis of comparatives involving multidimensional adjectives. Indeed, this is what I will suggest, which is rather different from the standard analysis.

Let us assume, for the sake of argument, that there are only two properties/dimensions associated with being clever: an ability to manipulate numbers, and an ability to manipulate people. One natural idea (e.g. Klein, 1980) is then to claim that John is cleverer than Mary iff John is better *both* in manipulating numbers *and* in manipulating people. More in general, let us assume that if '$>_C$' depends on '$>_{C_1}$', '$>_{C_2}$', and '$>_{C_3}$', then $x >_C y$ iff $x >_{C_1} y$, $x >_{C_2} y$, *and* $x >_{C_3} y$. Of course, this only gives rise to a strict partial order.[29] To come back to our example, suppose Sue is worse than John in manipulating numbers but better in manipulating people. According to the above suggested analysis, it is neither the case that John is cleverer than Sue, nor that Sue is cleverer than John, without them being necessarily equally clever. This is a typical possibility if 'cleverer than' only gives rise to a strict partial order. But perhaps we want to end up with more. I am not sure whether we always want more, but sometimes we certainly do.[30] If Sue is only a slightly bit worse than John in manipulating numbers, but much better in manipulating people, I take Sue to be cleverer than John. Or take for instance the adjectives 'useful' or 'beautiful'. It is obvious that one action, object, or individual can be more useful or beautiful than another and that the usefulness or beauty of something depends on various factors. Still, it is widely agreed in economics (and, I think, daily life) that 'being more useful than' gives rise to (at least) a strict weak

paratives with measure phrases.

[29]Recall that R gives rise to a strict partial order if it is irreflexive and transitive. For R to give rise to a weak order, it also has to be almost connected (cf. section 2).

[30]Lewis (1973), for instance, argues that comparative similarity between worlds is a weak order, rather than a partial order. For discussion, see Morreau (ms).

order. Similarly for constructions like 'better than', 'more beautiful than' and many others.

Let us assume that '$>_C$' depends on '$>_{C_1}$', '$>_{C_2}$', and '$>_{C_3}$'. How can we determine '$>_C$', such that this gives rise to a (strict) weak order rather than just a strict partial order? One idea would be to construct a lexicographical ordering, meaning that we first look only at '$>_{C_1}$', and only in case of a tie look at '$>_{C_2}$'. If x and y are then still indistinguishable, finally look at '$>_{C_3}$'. Lexicographical orderings, however, are the exception, rather than the rule. Can we not use a more 'democratic' procedure? We can, and the most natural way to reach this goal is to allow for interadjective comparisons. One way to go would be to follow Bale (2008) again: construct a 'universal' measurement system, and somehow construct '$>_C$' out of that. This is certainly possible, but the use a ratio-scale is again not required. Suppose, for instance, that we assume (only) co-ordinal invariance, that is, f and g (functions from $X \times \{C_1, C_2, C_3\}$ to $\mathbf{R}$) are considered to be 'the same' iff they can be related to each other by a strictly increasing monotone function.

Let us also assume that no-one of $\{>_{C_1}, >_{C_2}, >_{C_3}\}$ 'dictates' the ordering relation $>_C$ (for instance, we don't want '$>_C$' to be a lexicographical ordering based on $\{>_{C_1}, >_{C_2}, >_{C_3}\}$).[31] Now one can prove that under some natural assumptions $x \geq_C y$ iff $min_i f(\langle x, C_i\rangle) \geq min_i f(\langle y, C_i\rangle)$.[32] Alternatively, one could propose that $x \geq_C y$ iff $max_i f(\langle x, C_i\rangle) \geq max_i f(\langle y, C_i\rangle)$. Of course, we can make stronger assumptions on how to compare the different aspects of C-ness. If we assume that the mapping g from $X \times \{C_1, C_2, C_3\}$ to $\mathbf{R}$ is considered to be 'the same' as mapping f iff $\forall\langle x, C_i\rangle \in X \times \{C_1, C_2, C_3\}$: $g(\langle x, C_i\rangle) = \alpha \times f(\langle x, C_i\rangle) + \beta$, for some $\alpha, \beta > 0$, and moreover assign weighs to the aspects, we can end up with a 'C-er than' relation either defined as (i) $x \geq_C y$ iff $\sum_i w_i \times f(\langle x, C_i\rangle) \geq \sum_i w_i \times f(\langle x, C_i\rangle)$,[33] or as (ii) $x \geq_C y$ iff $\prod_i w_i \times f(\langle x, C_i\rangle) \geq \prod_i w_i \times f(\langle x, C_i\rangle)$. It is sometimes argued that if C depends on multiple dimensions, comparatives of the form 'x is C-er than y' always involve just one dimension, and the ordering is still just a weak ordering (or more). Though observing the frequent use of abstract comparatives like 'better than' or 'more useful than' makes this strong claim rather dubious, it cannot be denied that in practice we only take a limited set of dimensions into account when evaluating such a comparative. In fact, this can be easily modeled by making the weigh-function context dependent: $w_i = 0$, if dimension C_i is not

[31] This can be formalized by a notion of *permutation invariance*: Let π be a permutation of $\{C_1, C_2, C_3\}$. If $f(\langle x, C_i\rangle) = g(\langle x, \pi(C_i)\rangle)$ for all x and i, then $f(\langle y, C\rangle) > f(\langle z, C\rangle)$ iff $g(\langle y, C\rangle) > g(\langle z, C\rangle)$ for all $y, z \in X$.

[32] This is really the counterpart of Rawls's (1971) welfare function in political economy.

[33] This is the counterpart of Harsanyi's (1955) welfare function. In fact, Harsayni does not even use a weigh function.

taken into account. The main point of this section was to show that once we allow for interadjective measurement, comparisons involving multidimensional adjectives can be strengthened to weak orders in a natural way.

6 Conclusion

Thinking of semantics as the representation of linguistic meaning in terms of mathematical structures that we understand well, the main goal of this paper was to investigate what is required in order to account for several types of interadjective comparative statements. According to Stevens (1959) and Carnap (1966), in science one generally starts with classificatory concepts, later on invent orderings, which finally evolve to more and more fine-grained measurement scales. In this paper I discussed the relations between these three 'levels' of talking about phenomena. I indicated what kind of comparative statements can be accounted for by taking the classificatory concepts as basic, and discussed what more has to be assumed if we want to make sense of other comparative statements. I concentrated on interadjective comparisons and pointed out an interesting relation between this and the analysis of intersubjective comparisons of utility. I discussed a number of constructions that can be used to build interadjective comparisons, and I pointed out that such comparisons are required as well if one wants to account for the intuition that at at least some multidimensional adjectives give rise to a weak ordering relation. This paper does not solve the empirical question what representational structures, or constructions, we actually use. But I argued that some of the prominent analyses are less obvious than they might seem, and pointed to some other possibilities. I hope this was an (even?) more useful thing to do than it was clever.

References

[29] Arrow, K. (1950), 'A difficulty in the concept of social welfare', *Journal of Political Economy*, **58**: 328-346.

[29] Bale, A. (2008), 'A universal scale of comparison', *Linguistics and Philosophy*, **31**: 1-55.

[29] Benthem, J. van (1982), 'Later than late: on the logical origin of the temporal order', *Pacific Philosophical Quarterly* **63**: 193-203.

[29] Bierwisch, M. (1984), 'The semantics of gradation', in Bierwisch, M. and E. Lang (eds.), *Dimensional Adjectives*, pp. 71-262, Springer Verlag.

[29] Carnap, R. (1966), *Philosophical Foundations of Physics: An Introduction to Philosophy of Science*, Basic Books, New York.

[29] Doetjes, J. (2009), 'Incommensurability', in M. Aloni & K. Schulz (eds.), *Preproceedings A'dam Colloquium*, pp. 227-236.

[29] Harsanyi, J.C. (1955), 'Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility', *Journal of Political Economics*, **63**: 309-321.

[29] Kennedy, C. (1999), *Projecting the Adjective: The Syntax and Semantics of Gradability and Comparison*. Garland Press, New York.

[29] Klein, E. (1980), 'The semantics of positive and comparative adjectives', *Linguistics and Philosophy*, **4**: 1-45.

[29] Klein, E. (1991), 'Comparatives', in von Stechow, A. and D. Wunderlich (eds.), *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung*, Walter de Gruyter.

[29] Krantz, D., R. Duncan Luce, P. Suppes, and T. Tversky (1971), *Foundations of Measurement, Vol. 1*, Academic Press, New York and London.

[29] Lewis, D. (1973), *Counterfactuals*, Harvard University Press, Cambridge MA.

[29] McConnell-Ginet, S. (1973), *Comparative Constructions in English: A Syntactic and Semantic Analysis*, Phd thesis, University of Rochester.

[29] Morreau, M. (ms), 'It simply does not add up: trouble with overall similarity', manuscript.

[29] Rawls, J. (1971), *A Theory of Justice*, London: Harvard University Press.

[29] Rooij, R. van (2008), 'Comparatives and Quantifiers'. in Bonani and Hofherr (eds), *Empirical Issues in Syntax and Semantics 7*, pp..

[29] Rooij, R. van (2009), 'Revealed preference and satisficing behavior', in G. Bonano et. al. (eds.), Proceedings of LOFT 2008, Amsterdam University Press.

[29] Rooij, R. van (2010a), 'Explicit versus implicit comparatives', In P. Egre and N. Klinedinst *Vagueness and Language Use*, Palgrave MacMillan, New York.

[29] Rooij, R. van (2010b), 'Vagueness and linguistics', in Ronzitti (ed.), *Vagueness: A Guide*, Springer.

[29] Sassoon, G. (2010), 'Measurement theory in linguistics', *Synthese*, **174**: 151-180.

[29] Scott, D. and P. Suppes (1958), 'Foundational aspects of theories of measurement', *Journal of Symbolic Logic*, **23**: 113-128.

[29] Sen, A. K. (1970), *Collective Choice and Social Welfare*, North Holland, Amsterdam.

[29] Seuren, P. (1984), 'The comparative revisited', *Journal of Semantics*, **3**: 109-141.

[29] Shephard, R. (1962), 'The analysis of proximities: multidimensional scaling with an unknown distance function, I', *Psychometrika*, **27**, 125-140.

[29] Stechow, A. von (1984), 'Comparing semantic theories of comparison', *Journal of Semantics*, **3**: 1-77.

[29] Stevens, S. (1946), 'On the theory of scales of measurement', *Science*, **103**: 677-680.

[29] Stevens, S. (1959), 'Measurement, psychophysics, and utility', in C. W. Chungman & P. Ratoosh (eds.), *Measurement: Definitions and Theories*, Wiley, New York, pp. 129-143.

[29] Suppes, P. and J.L. Zinnes (eds.), *Basic Measurement Theory*, Handbook of Mathematical Psychology, Willey, New York.

[29] Tversky, A, (1977), 'Features of similarity', *Psychological Review*, **84**: 327-352.