In: L. Rudolph (Ed). Qualitative Mathematics for the Social Sciences: Mathematical Models for Research on Cultural Dynamics (pp. 105-136). London: Routledge.

1

The Sorites Paradox: A Behavioral Approach

Ehtibar N. Dzhafarov and Damir D. Dzhafarov

The issues discussed in this chapter can be traced back to the Greek philosopher Eubulides of Miletus. He lived in the 4th century BCE, a contemporary of Aristotle whom he, according to Diogenes Laërtius, "was constantly attacking" (Yonge, 1901, pp. 77–78). Eubulides belonged to what is known as the Megarian school of philosophy, founded by a pupil of Socrates named Euclid(es). Besides his quarrels with Aristotle we know from Diogenes Laërtius that Eubulides was the target of an epigram referring to his "false arrogant speeches", and that he "handed down a great many arguments in dialectics", mostly trivial sophisms of the kind ridiculed by Socrates in Plato's Dialogues. For example, the Horned Man argument asks you to agree that 'whatever you haven't lost you have', and points out that then you must have horns since you have not lost them.

Two of the "arguments in dialectics" ascribed to Eubulides, however, are among the most perplexing and solution-resistant puzzles in history. The first is the Liar paradox, which demonstrates the impossibility of assigning a truth value to the statement 'This statement is false'.* Eubulides's second 'serious' paradox, the Heap, is the subject of this chapter. It can be stated as follows. (1) A single grain of sand does not form a heap, but many grains (say $1,000,000$) do. (2) If one has a heap of sand, then it will remain a heap if one

---

Keywords: closeness, comparison, connectedness, matching, regular mediality, regular minimality, sorites, supervenience on stimuli, tolerant responses, vagueness.

*[See Chapter 2,p. 43*ff*, for more on the Liar. (Eds.)]

removes a single grain from it. (3) But, by removing from a heap of sand one grain at a time sufficiently many times, one can eventually be left with too few grains to form a heap. This argument is traditionally referred to by the name 'sorites' (from the Greek *σωρός* [*soros*] meaning 'heap'), with the adjective 'soritical' used to indicate anything 'sorites-related'. Thus, the Bald Man paradox which Diogenes Laërtius lists as yet another argument of Eubulides in dialectics is a 'soritical argument', because it follows the logic of the sorites but applies it to the example of the number of hairs forming or not forming a full head of hair.

Two Varieties of Sorites

This chapter is based on Dzhafarov & Dzhafarov (2010a,b), in which we proposed to treat sorites as a behavioral issue, with 'behavior' broadly understood as the relationship between stimuli acting upon a system (the 'system' being a human observer, a digital scale, a set of rules, or anything whatever). We present here a sketch of this treatment, omitting some of the more delicate philosophical points. Examples of behavioral questions pertaining to sorites include: Can a person consistently respond by different characterizations, such as 'is 2 meters long' and 'is not 2 meters long', to visually presented line segments a and b which only differ by one billionth of one percent? Is the person bound to say that these segments, a and b, look 'the same' when they are presented as a pair? But soritical questions can also be directed at non-sentient systems: Can a crude two-pan balance at equilibrium be upset by adding to one of the pans a single atom? Can the probability that this balance will remain at equilibrium change as a result of adding to one of the pans a single atom?

Sorites, viewed behaviorally, entails two different varieties of problems. The first, *classificatory sorites*, is about the identity or nonidentity of responses, or some properties thereof, to stimuli that are ' almost identical', 'differ only microscopically'. The second, *comparative sorites*, concerns 'match/not match'-type responses to *pairs of stimuli*, or more generally, response properties interpretable as indicating whether the two stimuli in a pair 'match' or 'do not match'. A prototypical example would be visually presented pairs of line segments with 'matching' understood as 'appearing the same in length'. Perhaps surprisingly, the two varieties of sorites turn out to be very different. The classificatory sorites is a logical impossibility, and our contribution to its analysis consists in demonstrating this impossibility by explicating its underlying assumptions on arguably the highest possible level of generality (using the mathematical language of Maurice Fréchet's proto-topological V-spaces). The comparative sorites (also called 'observational' in the philosophical literature) is, by contrast, perfectly possible: one can construct abstract and even physically realizable systems which exhibit soritical behavior of the comparative variety. This, however, by no means *has to be* the case for any system with 'match/not match'-type response properties. Most notably, we will argue that contrary to the widespread view this is not the case for the human comparative judgments, where comparative sorites contradicts a certain regularity principle supported by all available empirical evidence, as well as by the practice and language of the empirical research dealing with perceptual matching.

The compelling nature of the view that the human comparative judgments are essentially soritical is apparent in the following quotation from R. Duncan Luce (who used this view to motivate the introduction of the important algebraic notion of a semiorder).

> It is certainly well known from psychophysics that if "preference" is taken to mean which of two weights a person believes to be heavier after hefting them, and if "adjacent" weights are properly chosen, say a gram difference in a total weight of many grams, then a subject will be indifferent between any two "adjacent" weights. If indifference were transitive, then he would be unable to detect any weight differences, however great, which is patently false. (Luce, 1956, p. 179)

One way of conceptualizing this quotation so that it appears to describe a 'paradox' is this. (1) If the two weights being hefted and compared are the same, x and x, they 'obviously' match perceptually. (2) If one adds to one of the two weights a 'microscopic' amount ε (say, the weight of a single atom), the human's response to x and $x + \varepsilon$ cannot be different from that to x and x, whence x and $x + \varepsilon$ must still match perceptually. (3) But by adding ε to one of the weights many times one can certainly obtain a pair of weights x and $x + n\varepsilon$ that are clearly different perceptually.

This reasoning follows the logic of the *classificatory* sorites, and is thus logically invalid. This reasoning, however, seems intuitively compelling, which explains both why the philosophers call the comparative ('observational') sorites a sorites, and why psychophysicists need to deal with the classificatory sorites even if they are interested primarily in human comparative judgments.

Luce (1956) definitely did not imply a connection to the classificatory sorites of Eubulides. Rather, he simply presented as "well known" the impossibility of telling apart very close stimuli. This being a logically tenable position, we will see that the "well known" fact in question is in reality a theoretical belief not founded in empirical evidence. Almost everything in it contradicts or oversimplifies what we know from modern psychophysics. Judgments like 'x weighs the same as y' or 'x is heavier than y' given by human observers in response to stimulus pairs cannot generally be considered predicates on the set of stimulus pairs, as these responses are not uniquely determined by these stimulus pairs: an indirect approach is needed to define a matching relation based on these inconsistent judgments. When properly defined, the view represented by the quotation from Luce's paper loses its appearance of self-evidence.

Classificatory Sorites

The classificatory sorites is conceptually simpler than the comparative one and admits a less technical formal analysis. Within the framework of the behavioral approach we view the elements that the argument is concerned with (such as collections of grains of sand) as stimuli 'acting' upon a system and 'evoking' its responses. Thus, the stimuli may be electric currents passing through a digital ammeter which responds by displaying a number on its indicator; or the stimuli may be schematic drawings of faces visually presented to a human observer who responds by saying that the face is 'nice' or 'not nice'; or the stimuli may be appropriately measured weather conditions in May to which a flock of birds reacts by either migrating north or not. This is a very general framework which many examples can be molded to fit. In Eubulides's original argument, the stimuli are collections of sand grains, presented visually or described verbally, and the system responding by either 'form(s) a heap' or 'do(es) not form a heap' may be a human observer, if one is interested in factual classificatory behavior, or a system of linguistic rules, if one is interested in the normative use of language.

Supervenience, Tolerance, and Connectedness

To get our analysis off the ground, we would like to identify some properties which characterize stimulus-effect systems amenable to (classificatory) soritical arguments. Consider the following.

Supervenience assumption (Sup). There is a certain property of the system's responses to stimuli that—all else being equal—cannot have different values for different instances (replications) of one and the same stimulus. That is, there is a function π such that a certain property of the response of the system to stimulus x is $\pi(x)$. We call π the *stimulus-effect function*, and its values *stimulus effects*.[1]

Tolerance assumption (Tol). The stimulus-effect function $\pi(x)$ is 'tolerant to microscopic changes' in stimuli: if $x' \neq x$ is chosen sufficiently close to x, then $\pi(x') = \pi(x)$.

Connectedness assumption (Con). The stimulus set S contains at least one pair of stimuli a, b with $\pi(a) \neq \pi(b)$ such that one can find a finite chain of stimuli $a = x_1, \ldots, x_i, x_{i+1}, \ldots, x_n = b$ leading from a to b 'by microscopic steps': x_{i+1} is arbitrarily or maximally close to but different from x_i for $i = 1, \ldots, n-1$.

We have, of course, yet to define what precisely we mean by 'closeness' and 'connectedness by microscopic steps' here, but deferring that for the moment, it is not difficult to see that the conjunction Sup $\wedge$ Tol $\wedge$ Con is sufficient and necessary for formulating the classificatory sorites 'paradox'.

Classificatory Sorites. There exists a stimulus-effect system satisfying Sup, Tol, and Con.

It is clear that this statement is false: the three assumptions in questions are mutually inconsistent. Indeed, by Sup and Con we can fix a pair of stimuli a, b with $\pi(a) \neq \pi(b)$, connectable by a *classificatory soritical sequence* $x_1, \ldots, x_n$ with $a = x_1$, $b = x_n$, and x_{i+1} only 'microscopically' different from x_i for each i. By Tol, $\pi(x_i) = \pi(x_{i+1})$

[1] The response itself, e.g., 'heap' or 'not heap', may be viewed as a response property (namely, its identity or content), and this property may or may not be a stimulus effect. Other candidates for being stimulus effects can be such response properties as response time, response probability, the probability with which response time falls within a certain interval, etc.

for $i = 1, \ldots, n-1$, whence $\pi(a) = \pi(b)$, a contradiction. Therein lies the classificatory sorites.

Notice that if Sup is not satisfied, then Tol and Con simply cannot be formulated as above, as these formulations make use of a stimulus-effect function $\pi(x)$. Once Sup is accepted, simple mathematical examples can be constructed to witness the independence of Tol and Con. It is natural to ask whether Tol and Con can be formulated without an explicit reference to Sup, but this can readily be seen not to be an option, at least not without making Tol 'automatically' false. Indeed, if a property π_x of a response to stimulus x is not a function of x, then π_x will generally be different from π_y even if y is a replication of x, let alone close to but different from x.

The fact that Sup is indispensable for the formulability of the classificatory sorites leads one to reject the philosophical tradition of relating soritical issues to 'vague predicates'. A vague predicate is defined as one whose truth value is not determinable at least for some objects to which the predicate applies, whether the structure of its truth values is dichotomous (true, false), trichotomous (true, false, not known), or the entire interval between 0 (false) and 1 (true). Here we see a major advantage of the behavioral approach to sorites. One may very well argue about the truth value structure of the predicate 'form(s) a heap', and one may suggest that for certain values of x this predicate's truth value is indeterminate when applied to a collection of x grains of sand. However, a statement like "in this trial, this observer responded to x grains of sand by saying 'they form a heap'" or "in this trial, this observer did not produce a response to x grains of sand when asked to choose between 'they form a heap' and 'they do not form a heap'" is true or false in the simplest sense, with no controversy involved. The 'observer' in this examples can very well be replaced with a set of linguistic rules or the group of expert language users if one is interested in the normative rather than factual use of language.[2]

From the behavioral point of view (and in agreement with the position, stated with admirable clarity in the dictionary article by Peirce, 1901/1960), a 'vague predicate' is merely a special case of an inconsistent response. If we call the predicate 'form(s) a heap' vague, this is because the choice of an allowable response associated

[2] In the philosophical literature, Varzi (2003) comes close to the behavioral approach by arguing that soritical issues are essentially non-semantic and are not confined to linguistic phenomena.

with this predicate (such as 'yes, it is true', 'possibly', or 'I don't know') is not uniquely determined by the number x of the sand grains to which it applies. Thus, a human observer is likely to classify one and the same collection of x grains sometimes as forming and sometimes as not forming a heap (and sometimes neither if this is an option); and in a group of competent speakers of the language some will choose one response, and some another. This implies a violation of Sup and the impossibility of formulating the soritical argument with this predicate. The behavioral scientist in a situation like this would likely redefine the stimulus-effect function $\pi = \pi(x)$ as the probability distribution on the set of all allowable responses, the hypothesis being that this probability distribution is now uniquely determined by stimuli. Thus, if the allowable responses are 'form(s) a heap' and 'do(es) not form a heap', then the hypothesis is that for some probability function $p(x)$, called a *psychometric function* in psychophysics,

$$\pi(x) = p(x) = \Pr[\text{'a collection of } x \text{ grains of sand is a heap'}].$$

Of course, to say that a probability p of a response to x is an effect of the stimulus x amounts to treating probabilities as occurring at individual instances of x 'within' the system responding to x, rather than characterizing patterns of the system's behaviors over a potential infinity of instances of x. While this view may encounter philosophical misgivings, it is routine in the established conceptual schemes of probability theory, physics, and behavioral sciences. Our analysis is not critically based on accepting this 'probabilistic realism', but the class of physically realizable response properties uniquely determined by stimuli may get precariously small if one rejects it.[3] Without allowing probability distributions over responses to function as legitimate stimulus effects one would often have to declare sorites altogether unformulable and hence automatically dissolved, or would have to seek additional factors to include in the description of stimuli.

There are a number of avenues for redefining stimuli in order to achieve the compliance of some response property (deterministic or probabilistic) with Sup. Thus, one might think it important to take into account sequential effects, that is, to make the response property in question dependent on a sequence of previously presented

[3] In the context of the comparative sorites (p. 13 *ff*), Hardin (1988) argued for the necessity of taking into account the probabilistic nature of responses to stimuli.

stimuli, or even on both the previous stimuli and the responses given to each of them. In either case we deal with some form of compound stimuli, the space of which we can endow (not necessarily in a unique way) with a closeness structure based on that of the original space of stimuli. For instance, if in the sequence $x_0, x_1, x_2, \ldots$ of stimuli each x_{i+1} differs from x_i 'microscopically', then the same can be said of x^*_{i+1} and x^*_i in the sequence $x^*_0, x^*_1, x^*_2, \ldots$ of compound stimuli $x^*_i = \{x_j\}_{j \leq i}$. If now a stimulus-effect function π, such as the probability of saying 'form(s) a heap', is uniquely determined by such finite sequences of successive stimuli, the characterizations Sup, Tol, and Con will be formulable for this function on the redefined stimulus set, and our analysis applies with no modifications.

To see the generality of our approach, one may even consider a radical redefinition of stimuli (by no means feasible for scientific purposes) which consists in taking stimulus instances as part of stimulus identities, so that each stimulus is formally characterized by a pair (x, t) where x is the stimulus's physical value, and t designates an 'instance' at which it occurs, say, a trial number. With this redefinition no stimulus (x, t) is replicable, because of which every response to (x, t) can be viewed as a function of (x, t), a stimulus effect. Given an 'initial' closeness measure between stimulus values x and y, and the conventional distance $|t - t'|$ between time moments t and t', it is easy to see that Tol in this situation means that (x, t) and (y, t') evoke identical responses if x and y are sufficiently close and $|t - t'|$ sufficiently small; and Con means that for some (a, t) and (b, t') which evoke different responses one can find a sequence $(a, t) = (x_1, t_1), \ldots, (x_n, t_n) = (b, t')$ whose successive elements fall within the sphere of Tol. As Sup here is satisfied 'automatically', our analysis again applies with no modifications.

Closeness and connectedness

It is apparent from our formulation of Tol and Con that we need conceptual means of saying that two distinct stimuli x and y can be chosen 'as close as one wishes' or 'as close as possible'. The meaning is clear in the case the stimulus set is endowed with a metric. Thus, if stimulus values (say, time intervals) are represented by real numbers, y can be chosen arbitrarily close to x by making the difference $|x - y|$ arbitrarily small. If, as it is the case with grains of sand, stimulus values are represented by integers, y is as close as possible to x (without being equal to it) if $|x - y| = 1$. The requirement of a full-fledged metric, however, is too stringent in general, and it is moreover unnecessary. The analysis of the classificatory sorites can be carried out at a much higher (arguably, the highest possible) level of generality using the following concept due to Fréchet (1918).

Definition 1. A *V-space* on a nonempty set S is a pair $\{S, \{\mathcal{V}_x\}_{x\in S}\}$ where $\mathcal{V}_x$, for each $x \in S$, is a collection of subsets of S satisfying (1) $\mathcal{V}_x \neq \varnothing$, (2) if $V \in \mathcal{V}_x$ then $x \in V$. For each $x \in S$, any element V of $\mathcal{V}_x$ is called a *vicinity* of x. Any set of vicinities obtained by choosing one element of $\mathcal{V}_x$ for every $x \in S$ is called a *V-cover* of S.

For each $x \in S$, each $V \in \mathcal{V}_x$ represents the stimuli which are close to x *in some sense*, namely, in the sense of belonging to V. In particular, since x belongs to each of its vicinities, x is 'close' to itself in all possible senses. (Fig. 1.1 illustrates 'closeness' in a V-space.)

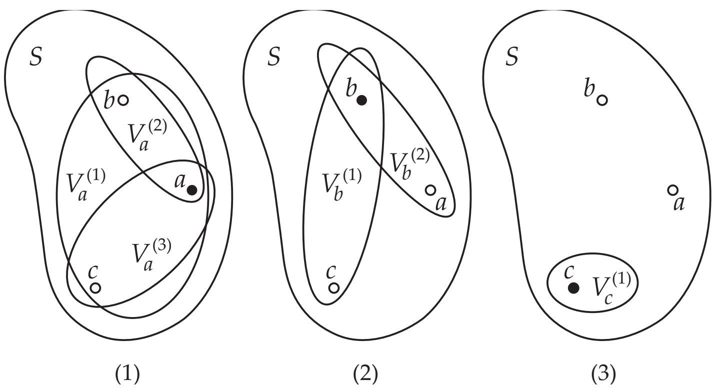


Figure 1.1: An example of a V-space $\{S, \{\mathcal{V}_x\}_{x\in S}\}$. The set S consists of all points within the large outlined area, including the points a, b and c shown. The vicinities $V_a^{(1)}, V_a^{(2)}, V_a^{(3)}$ of the point a, which together comprise $\mathcal{V}_a$, are indicated by the ovals in (1), and similarly for $\mathcal{V}_b = \{V_b^{(1)}, V_b^{(2)}\}$ in (2) and $\mathcal{V}_c = \{V_c^{(1)}\}$ in (3). These three sets of vicinities determine the closeness relations among a, b, and c. Thus, c is close to a in the sense of belonging to $V_a^{(1)}$ and $V_a^{(3)}$; c is close to b in the sense of belonging to $V_b^{(1)}$; and a and b are not close to c in any sense (*i.e.*, they are 'not close at all' to c) as they do not belong to any vicinity of c. Note that, at this level of generality, closeness is not a symmetric relation. Any V-cover of S will contain $V_c^{(1)}$, one and only one of $V_b^{(1)}, V_b^{(2)}$, and one and only one of $V_a^{(1)}, V_a^{(2)}, V_a^{(3)}$.

The notion of V-space obviously generalizes that of a topological space, and in particular, a V-space on any metric space (S, d) can be obtained by letting $\mathcal{V}_x$ consist of all open balls

$$B_x(\varepsilon) = \{u \in S\colon d(x,u) < \varepsilon\}$$

where $\varepsilon > 0$. In general, however, V-spaces provide for a notion of closeness that does not have to be numerical.

Next we need to use this notion of closeness to define connectedness. Eubulides's Heap riddle would not be perplexing were it not for the fact that by adding one grain of sand at a time one can obtain from a very small collection of grains of sand a collection large enough to form a heap. The key to defining connectedness in the general language of V-spaces is in replacing the notion of being connectable 'by microscopic steps' (which requires a quantitative measure of closeness) with the notion of being connectable, from each choice of vicinities covering the space, by a chain of sequentially intersecting vicinities.

Definition 2. A point $a \in S$ is *V-connected to* a point $b \in S$ in a V-space $\{S, \{\mathcal{V}_x\}_{x\in S}\}$ if for any V-cover $\{V_x\}_{x\in S}$ of S one can find a finite chain of points $x_1, x_2, \ldots x_{n-1}, x_n \in S$ such that (1) $a = x_1$, (2) $b = x_n$, (3) $V_{x_i} \cap V_{x_{i+1}} \neq \varnothing$ for $i = 1, \ldots n-1$. A V-space $\{S, \{\mathcal{V}_x\}_{x\in S}\}$ is *V-connected* if any two points in S are V-connected in $\{S, \{\mathcal{V}_x\}_{x\in S}\}$.

Consider the following example. Let $S = \mathbb{N}$ be the set of all natural numbers $0, 1, 2, \ldots$, and give it a V-space structure by defining, for each $n, k \in \mathbb{N}$, $V_{n,k} = \{n, n+1, \ldots, n+k\}$ and $\mathcal{V}_n = \{V_{n,k}\colon k > 0\}$. Then any two numbers $a < b$ in this space are V-connected because for any V-cover, the chain $x_1, x_2, \ldots, x_{b-a+1}$ with $x_i = a + i - 1$ satisfies properties (1)–(3) of Definition 2. If, however, we take instead $\mathcal{V}_n = \{V_{n,k}\colon k \geq 0\}$, then no two elements can be V-connected, as witnessed by the V-cover $\{V_{n,0}\colon n \in \mathbb{N}\}$.

It is easy to see that the relation of 'being connected to' is an equivalence relation, whence we immediately have the following lemma.

Lemma 1. *For any V-space $\{S, \{\mathcal{V}_x\}_{x\in S}\}$, the set S is a union $\bigcup S_\gamma$ of pairwise disjoint nonempty subsets (the* V-components *of S) such that any two points in every V-component are V-connected and no two points belonging to different V-components are V-connected.*

We also need the following definition to formalize Sup and Tol.

Definition 3. Given a V-space $\{S, \{\mathcal{V}_x\}_{x\in S}\}$ and an arbitrary set R, any function $\pi : S \to R$ is a *stimulus-effect function* (and R a set of *stimulus effects*). A stimulus-effect function π is called *tolerant at* $x \in S$ in $\{S, \{\mathcal{V}_x\}_{x\in S}\}$ if there is a vicinity $V_x \in \mathcal{V}_x$ on which π is constant; π is *tolerant* if it is tolerant at every point.

Thus, R can be the two-element set {'form(s) a heap','do(es) not form a heap'}; or the set $[0, 1]$ representing the probabilities of choosing the response 'form(s) a heap' over 'do(es) not form a heap'; or R can be the set of all probability distributions on $[0, 1]$, with $x \in [0, 1]$ representing the degree of confidence with which a stimulus is judged to form a heap.

Dissolving the classificatory sorites 'paradox'

The reward for formulating the soritical concepts on such a high level of generality is that the classificatory sorites 'paradox' can now be dissolved by means of a simple mathematical theorem.

Theorem 1 (No-tolerance theorem)**.** *Let $\{S, \{\mathcal{V}_x\}_{x\in S}\}$ be a V-space and $\pi : S \to R$ a stimulus-effect function, such that S contains two V-connected elements a, b for which $\pi(a) \neq \pi(b)$. Then π is not tolerant: there is at least one $x \in S$ such that π is nonconstant on any vicinity of x ('however small').*[4]

N.B. The words 'however small' are added for emphasis only and do not imply that the vicinities have numerical sizes.

Proof. Assume π is tolerant: every x has a vicinity V_x^* such that π is constant on V_x^*. The set $\{V_x^*\}_{x\in S}$ is a V-cover of S, and a, b being V-connected, one can form a sequence $V_{x_1}^*, \ldots, V_{x_n}^*$ satisfying (1)–(3) of Definition 2. Then, denoting by y_i an arbitrary element of $V_{x_i}^* \cap V_{x_{i+1}}^*$ for $i = 1, \ldots n-1$, we would have $\pi(x_i) = \pi(y_i) = \pi(y_{i+1}) = \pi(x_{i+1})$ (since $y_i, y_{i+1} \in V_{x_{i+1}}^*$), whence $\pi(a) = \pi(x_1) = \pi(x_n) = \pi(b)$, contradicting the premise $\pi(a) \neq \pi(b)$. □

[4]One may recognize in this formulation a generalized version of what is known as the epistemic dissolution of the classificatory sorites, proposed by Sorensen (1988a,b) and Williamson (1994, 2000), except that they apply this dissolution to vague predicates. This would only be legitimate if vague predicates were assigned to stimuli consistently, but then they would not be considered vague.

Corollary 1. *No nonconstant function on a V-connected V-space is tolerant.*

One can easily recognize in Theorem 1 the spelled-out version of the classical (classificatory) sorites, taken as a *reductio ad absurdum* proof of the incompatibility of Sup, Tol, and Con. More precisely, this incompatibility is formulated as $\mathsf{Sup} \wedge \mathsf{Con} \Longrightarrow \neg\mathsf{Tol}$. In fact, when discussing classificatory sorites, Sup and Con are almost always assumed implicitly, although Sup is sometimes mentioned as an innocuous premise. The 'paradox' (in these cases) is thus dissolved by pointing out the inescapable truth of the assertion our intuition often finds hard to accept:

Non-tolerance principle ($\neg\mathsf{Tol}$)**.** If Sup and Con hold, then there is at least one point $x_0 \in S$ in every vicinity of which ('however small'), the stimulus-effect function $\pi(x)$ is nonconstant.

Note that it may very well be that every single point in S complies with $\neg\mathsf{Tol}$, and this is probably the case in a host of situations with continuously varying stimulus effect (as a mathematical example, consider the identity function, mapping any stimulus into itself).

To prevent misunderstanding, $\neg\mathsf{Tol}$ does not imply that the responding system can be used to measure some of the stimulus values with absolute precision. The situation here is very much like the one with a stopped clock: it shows the correct time twice a day, but one cannot determine when. In order to distinguish a stimulus x from its arbitrarily close neighbors x' by means of a stimulus-effect function $\pi(x)$ a human researcher must know that x being presented on two different instances is indeed one and the same x, and that x' presented on another instance is not the same as x. This amounts to having a system identifying stimuli being presented, $\iota_S(x)$, and a system identifying the stimulus effects being recorded, $\iota_R(\pi(x))$, and hence to having yet another stimulus-effect function besides $\pi(x)$ whose values react to arbitrarily small differences from precisely the same stimulus x (something not impossible but definitely not deliberately construable unless the stimuli have been identified by some $\iota'_S(x)$, which assumption leads to an infinite regress).

For completeness, we should mention that there is another way of approaching the inconsistency of $\mathsf{Sup} \wedge \mathsf{Tol} \wedge \mathsf{Con}$, formulable as $\mathsf{Sup} \wedge \mathsf{Tol} \Longrightarrow \neg\mathsf{Con}$. We can define vicinities as constant-response areas of the stimulus set S; we call them 'pi-vicinities' since we denote the stimulus-effect function by π. This makes the stimulus-effect function 'automatically' tolerant, and one sees subsequently that no two points in S are V-connected unless they map into one and the same stimulus effect.

Definition 4. Given a nonempty set S, an arbitrary set R, and a stimulus-effect function $\pi : S \to R$, the *pi-vicinity* of $x \in S$ is the set P_x of all $x' \in S$ such that $\pi(x') = \pi(x)$. The pair $\{S, \{P_x\}_{x \in S}\}$ is called the *pi-space associated to* π.

Lemma 2. *Any pi-space $\{S, \{P_x\}_{x \in S}\}$ uniquely corresponds to the V-space on S in which the only vicinity of $x \in S$ is P_x. The collection of the sets $\{P_x\}_{x \in S}$ is the only V-cover of S in this V-space.*

Proof. It is clear that $\{S, \{\mathcal{V}_x\}_{x \in S}\}$ with $\mathcal{V}_x = \{P_x\}$ satisfies Definition 1. □

Lemma 3. (1) *The pi-space $\{S, \{P_x\}_{x \in S}\}$ associated to $\pi : S \to R$ is uniquely determined by π.* (2) *π is tolerant in the corresponding V-space $\{S, \{\mathcal{V}_x = \{P_x\}\}_{x \in S}\}$.*

Proof. Immediate consequences of Definition 4, Lemma 2, and Definition 3. □

This yields the following alternative dissolution of the classificatory sorites.

Theorem 2 (No-connectedness theorem)**.** *Given a stimulus-effect function $\pi : S \to R$ and its associated pi-space $\{S, \{P_x\}_{x \in S}\}$, elements $a, b \in S$ are V-connected in the corresponding V-space $\{S, \{\mathcal{V}_x = \{P_x\}\}_{x \in S}\}$ if and only if $\pi(a) = \pi(b)$.*

Proof. An immediate consequence of Definition 2 and the fact that either $P_x = P_y$ or $P_x \cap P_y = \varnothing$, for any $x, y \in S$. □

Comparative Sorites

The *comparative* sorites pertains to situations in which a system responds to *pairs* of stimuli (x, y), and there is some binary response property which is uniquely determined by (x, y) and whose values are interpretable as two complementary relations, 'x is matched by y' and 'x is not matched by y'. A prototypical example would be an experiment in each trial of which a human observer is shown a pair of line segments and asked whether they are or are not of the same length, or two circles of light and asked which of them is brighter. The matching relation in both cases is computed from probabilities of the observer's answers, as discussed later. Thus, we regard the comparative sorites as pertaining to what in psychophysics would be called *discrimination* or *pairwise comparison* tasks.

Comparing stimuli

The comparative sorites 'paradox' can be intuitively stated as follows.

Comparative Sorites. A set of stimuli S acting upon a system and presentable in pairs may contain a finite sequence of stimuli $x_1, \ldots, x_n$ such that 'from the system's point of view' x_i is matched by x_{i+1} for $i = 1, \ldots, n-1$, but x_1 is not matched by x_n.

We call a sequence $x_1, \ldots, x_n$ as above a *comparative soritical sequence*.

At the outset, the comparative sorites may seem very similar to the classificatory sorites, and it is natural to ask whether the former may be only a special case of the latter. This turns out not to be the case. It is true that nothing prevents one from redefining a pair of stimuli (x_i, x_j) into a single 'bipartite' stimulus x_{ij}, and treating 'match' and 'not match' as classificatory responses to x_{ij}. However, given a sequence $x_{12}, x_{23}, \ldots, x_{n-1,n}, x_{1n}$, we can only apply to it the rationale of the classificatory sorites provided each term in this sequence is only 'microscopically' different from its successor. While this may be the case for $x_{i,i+1}$ and $x_{i+1,i+2}$ for each $i = 1, \ldots, n-2$ (perhaps because x_i is very close to x_{i+1}, which in turn is very close to x_{i+2}), there is in general no reason at all to think that $x_{n-1,n}$ is only 'microscopically' different from x_{1n}, assuming a reasonable notion of closeness between the two can be formulated at all.[5] Thus,

[5]Nor can one consider the classificatory sorites a special case of the comparative one. Given a purported classificatory soritical sequence $x_1, \ldots, x_n$ under some stimulus-effect function π, we can indeed recast $\pi(x_i)$ into a function $f(x_i, x_i + 1)$ loosely interpretable as a 'comparison' of x_i with x_{i+1} for each $i = 1, \ldots, n-1$. But then the logic of the classificatory sorites would lead to the 'comparison' of x_{n-1} with x_n being different from that of x_1 with x_2 (while in the comparative sorites the two pairs produce the same effect, 'match'), with nothing in this logic necessitating

the existence of comparative soritical sequences is not automatically precluded by the analysis of the classificatory sorites given in the previous section. Indeed, it is easy to construct simple mathematical examples of such sequences, arguably the simplest of them being as follows.

Example 1. Fix $\varepsilon > 0$. Let the stimulus set S be the set $\mathbb{R}$ of reals, and define

$$\text{`}y\text{ matches }x\text{'} \iff |y - x| \leq \varepsilon.$$

Then any sequence $0, \delta, 2\delta$ with $\varepsilon/2 < \delta \leq \varepsilon$ is a comparative soritical sequence.

This example also precludes the possibility of reducing the comparative sorites to the classificatory one by postulating the existence of a stimulus-effect function $\pi(x)$ such that

$$\text{`}x\text{ is matched by }y\text{'} \iff \pi(x) = \pi(y).$$

If such a function could always be found, the comparative sorites would indeed be obtained as a 'logical consequence' of the classificatory one and would then be ruled out together with the latter. In the preceding example, however, it is readily seen that given any non-constant stimulus-effect function π, there exists an x with $\pi(x) \neq \pi(y)$ for some y with $x < y \leq x + \varepsilon$, even though $|y - x| \leq \varepsilon$ and so 'y matches x'.

We see that, on the one hand, the comparative sorites cannot be ruled out as a logical inconsistency, as it was in the case of the classificatory one. On the other hand, the comparative sorites 'paradox' is rarely if ever presented as an exercise in constructing abstract mathematical examples like the one above, and is generally assumed to apply to systems which, in their responses, resemble human comparative judgments. That human comparative judgments are soritical is often considered self-evident and well-known. As we shall see in the next subsection, this assumption contradicts a certain psychophysical principle (Regular Mediality/Minimality) proposed for comparative judgments. In systems similar to human comparative judgments, for which it is plausible that this principle holds, the 'matching' relation must satisfy (a certain form of) transitivity, thus

x_1 to be 'compared' with x_n (while in the comparative sorites this comparison is critical).

compelling the system to behave more in accordance with the following example than with Example 1.

Example 2. Fix $\lambda > 0$. Let the stimulus set be $S = \mathbb{R}_{\geq 0}$ (nonnegative reals), and define

$$\text{`}y \text{ matches } x\text{'} \iff \lfloor \lambda x \rfloor = \lfloor \lambda y \rfloor,$$

where $\lfloor a \rfloor$ denotes the floor of a, *i.e.*, the greatest integer $\leq a$. Then the relation 'matches' is reflexive, symmetric, and transitive, so no comparative soritical sequence involving this relation is possible.[6]

Stimulus areas

It is clear from the comparison of the two examples above that transitivity or lack thereof is at the heart of dealing with comparative soritical sequences. We will see, however, that with the recognition of the fact that two stimuli being compared belong to distinct 'stimulus areas' (the notion we explain next), the notion of transitivity should be approached with some caution, and the transitivity sometimes has to be formulated differently from the familiar, 'triadic' way (x is matched by y and y is matched by z, hence x is matched by z). For the same reason ('stimulus areas') the same degree of caution should be exercised in approaching the properties of reflexivity and symmetry, which most writers seem to take for granted.[7]

The notion of distinct 'stimulus areas' is both simple and fundamental. To say meaningfully that two physically identical stimuli, x and x, are judged as being the same or different, the two x's have to designate identical values of two otherwise different stimuli. If not for this fact we would have a single stimulus rather than two stimuli with identical values, and we would not be able to speak of pairwise comparisons. Thus, one of the two stimuli can be presented on the left and another on the right from a certain point, or one presented chronologically first and the other second. Stimuli, therefore,

[6] This agrees with the obvious fact that $\lfloor \lambda x \rfloor$ can be viewed as a stimulus-effect function defined on individual number-stimuli, so comparative soritical sequences here are ruled out by the nonexistence of classificatory ones.

[7] Thus, Goodman (1951/1997), Armstrong (1968), Dummett (1975), and Wright (1975) view it as obvious that perceptual matching is intransitive, Graff (2001); Jackson & Pinkerton (1973) argue for its transitivity (in the conventional, 'triadic' sense), and all of them consider it self-evident that perceptual matching is reflexive and symmetric.

should be referred to by both their *values* (for example, length) and their *stimulus areas* (for example, left and right).[8] The complete reference here is then (x, left) and (y, right), or more briefly, $x^{(l)}$ and $y^{(r)}$.

Stimulus areas need not be defined only by spatiotemporal positions of stimuli. Thus, two line segments compared in their length may be of two different fixed orientations, and two patches of light compared in their brightness may be of two different fixed colors (in addition to occupying different positions in space or time). Nor should there be only two distinct observation areas: pairs of light patches compared in brightness can appear in multiple pairs of distinct spatial positions, and can be of various colors.

The sets to which we shall address our formal analysis will consequently be of the form $S \times \Omega$, where S is a set of stimulus *values* and Ω a set of stimulus *areas*, both containing at least two elements. (We will continue to use the more convenient notation $x^{(\omega)}$ for the stimulus $(x, \omega) \in S \times \Omega$.) The most basic property of the matching relation M is then

$$x_1^{(\omega)}\,\mathrm{M}\,x_2^{(\omega')} \Longrightarrow \omega \neq \omega',$$

that is, that we do not compare stimulus values from the same stimulus area. This implies, in particular, that M is antireflexive:

$$\neg x^{(\omega)}\,\mathrm{M}\,x^{(\omega)}$$

holds for all $x^{(\omega)}$.[9]

For $\omega, \omega' \in \Omega$, the sets $S \times \{\omega\}$ and $S \times \{\omega'\}$ may simply be viewed as sets with different, further unanalyzable elements. We could, in fact, replace $S \times \Omega$ with $\{S_\omega\}_{\omega \in \Omega}$ treating thereby Ω as an indexing sets for a collection of sets otherwise unrelated to each other. This is an important point in view of situations where one would want to speak of matching between entities of different natures, *e.g.*, abilities of examinees and difficulties of the tests offered to them (as is routinely done in psychometric models). We prefer, however, an intermediate notational approach: we write $S \times \Omega$ as a

---

[8]The term used in Dzhafarov (2002), where the concept was introduced in a systematic fashion, was 'observation area', but 'stimulus area' seems preferable if the present analysis is also to apply to non-perceptual responses.

[9]This should not be confused with the statement that $\neg x^{(\omega)}\,\mathrm{M}\,x^{(\omega')}$ holds for all $x^{(\omega)}$ and $x^{(\omega')}$. We are perfectly free to compare two stimuli with the same value, so long as they belong to different stimulus areas, and these stimuli may match or not match.

reminder that stimuli in different stimulus areas may have identical values, but we treat $x^{(\omega)}$ 'holistically', saying, *e.g.*,

for any $a^{(\omega)}, b^{(\omega')} \in S \times \Omega$, the stimuli $a^{(\omega)}$ and $b^{(\omega')}$ are...

instead of

for any $\omega, \omega' \in \Omega$ and $a, b \in S$, the stimuli $a^{(\omega)}$ and $b^{(\omega')}$ are...

Psychophysics of matching

Let us begin with a situation involving only two stimulus areas, say,

$$\Omega = \{l, r\},$$

standing for 'left' and 'right'. In modern psychophysics (cf. Dzhafarov, 2002, 2003), matching relations on a given system $S \times \{l, r\}$ are defined from discrimination probability functions, as follows. For an observer asked to say whether two stimuli are the same or different, either with respect to a specified subjective property or overall, but ignoring the conspicuous difference in the stimulus areas, we can form a 'probability of being judged to be different' function,

$$\psi(x^{(l)}, y^{(r)}) = \Pr[x^{(l)} \text{ and } y^{(r)} \text{ are judged to be different}].$$

If the stimuli are compared with respect to a specified property, and if this property is linearly ordered (as in the cases of length, brightness, attractiveness, etc.), then the question can also be formulated in terms of which of the two stimuli has a greater amount of this property, and we can form a 'probability of being judged to be greater' function,

$$\gamma(x^{(l)}, y^{(r)}) = \Pr[y^{(r)} \text{ is judged to be greater than } x^{(l)}].$$

We can then define the 'matching' relation, henceforth denoted by M, either by

$$x^{(l)} \,\mathrm{M}\, y^{(r)} \text{ iff } \psi(x^{(l)}, y^{(r)}) = \min_z \psi(x^{(l)}, z^{(r)})$$
$$y^{(r)} \,\mathrm{M}\, x^{(l)} \text{ iff } \psi(x^{(l)}, y^{(r)}) = \min_z \psi(z^{(l)}, y^{(r)})$$

if dealing with ψ, or by

$$x^{(l)} \,\mathrm{M}\, y^{(r)} \text{ iff } \gamma(x^{(l)}, y^{(r)}) = 1/2$$
$$y^{(r)} \,\mathrm{M}\, x^{(l)} \text{ iff } \gamma(x^{(l)}, y^{(r)}) = 1/2$$

if dealing with γ. (No claim is being made that we get the same notion of matching in both cases.)

We now present two properties of the matching relation M that are critical for our treatment of comparative sorites. These properties, formulated and developed in Dzhafarov (2002, 2003) and Dzhafarov & Colonius (2006), constitute a principle which is called *Regular Mediality* or *Regular Minimality*, depending as it is applied to (a function like) γ or to ψ. These properties are formulated here in a form better suited to the present context of studying the comparative sorites, but their formulations in psychophysics are entirely unrelated to and unmotivated by soritical issues.

Regular Mediality/Minimality, Part 1 (RM1)**.** For every stimulus in either of the two stimulus areas, one can find a stimulus in the other stimulus area such that if $x^{(l)}$ and $y^{(r)}$ are the stimuli in question then $x^{(l)}\,\mathrm{M}\,y^{(r)}$ and $y^{(r)}\,\mathrm{M}\,x^{(l)}$.

To formulate the second property, we need the following notion. We call two stimuli in a given stimulus area *equivalent* if they match exactly the same stimuli in the other stimulus area. So, $x_1^{(l)}$ and $x_2^{(l)}$ are equivalent, in symbols $x_1^{(l)}\,\mathrm{E}\,x_2^{(l)}$, if

$$y^{(r)}\,\mathrm{M}\,x_1^{(l)} \Longleftrightarrow y^{(r)}\,\mathrm{M}\,x_2^{(l)}$$

for every $y^{(r)}$; and $y_1^{(r)}\,\mathrm{E}\,y_2^{(r)}$ if

$$x^{(l)}\,\mathrm{M}\,y_1^{(r)} \Longleftrightarrow x^{(l)}\,\mathrm{M}\,y_2^{(r)}$$

for every $x^{(l)}$.[10]

Regular Mediality/Minimality, Part 2 (RM2)**.** Two stimuli in one stimulus area are equivalent if there is a stimulus in the other area which matches both of them, *i.e.*,

$$\text{if } x_1^{(l)}\,\mathrm{M}\,y^{(r)} \text{ and } x_2^{(l)}\,\mathrm{M}\,y^{(r)} \text{ then } x_1^{(l)}\,\mathrm{E}\,x_2^{(l)},$$
$$\text{if } y_1^{(r)}\,\mathrm{M}\,x^{(l)} \text{ and } y_2^{(r)}\,\mathrm{M}\,x^{(l)} \text{ then } y_1^{(r)}\,\mathrm{E}\,y_2^{(r)}.$$

[10] In Dzhafarov & Colonius (2006) the equivalence is defined in a stronger way: $x_1^{(l)}\,\mathrm{E}\,x_2^{(l)}$ if $\psi(x_1^{(l)}, y^{(r)}) = \psi(x_2^{(l)}, y^{(r)})$ for all $y^{(r)}$, and $y_1^{(r)}\,\mathrm{E}\,y_2^{(r)}$ if $\psi(x^{(1)}, y_1^{(r)}) = \psi(x^{(1)}, y_2^{(r)})$, for all $x^{(1)}$.

In regard to γ, the relations $x^{(l)}\,\mathrm{M}\,y^{(r)}$ and $y^{(r)}\,\mathrm{M}\,x^{(l)}$ mean one and the same thing, $\gamma(x^{(l)}, y^{(r)}) = 1/2$. For γ, then, RM1 requires only that, for every $x_0^{(l)}$, the function $y \mapsto \gamma(x_0^{(l)}, y^{(r)})$ reaches the median level $1/2$ at some point $y_0^{(r)}$, and then it follows that the function $x \mapsto \gamma(x^{(l)}, y_0^{(r)})$ reaches the median level at $x_0^{(l)}$. In regard to ψ, however, RM1 is more restrictive, requiring not only that the functions $y \mapsto \psi(x_0^{(l)}, y^{(r)})$ and $x \mapsto \psi(x^{(l)}, y_0^{(r)})$ reach their minima at some points, but also that, for every x_0, if y_0 minimizes the function $y \mapsto \psi(x_0^{(l)}, y^{(r)})$ then x_0 minimizes the function $x \mapsto \psi(x^{(l)}, y_0^{(r)})$. Unlike with γ, RM1 imposes nontrivial restrictions on the properties of the function ψ (Kujala & Dzhafarov, 2008).

To understand RM2, note that it is satisfied trivially if all matches are determined uniquely, *i.e.*, if for each $y^{(r)}$ there is only one $x^{(l)}$ satisfying $x^{(l)}\,\mathrm{M}\,y^{(r)}$, and for each $x^{(l)}$ there is only one $y^{(r)}$ satisfying $y^{(r)}\,\mathrm{M}\,x^{(l)}$. However, this is not generally the case, and whether it is the case in specific cases depends on one's choice of the physical description of stimuli. The most familiar example is that of matching isoluminant colors. A given color on the right, $y^{(r)}$, usually matches a single color on the left, $x^{(l)}$, provided that $x^{(l)}$ is identified, say, by its CIE coordinates. But if $x^{(l)}$ is identified by its radiometric spectrum, then $y^{(r)}$ matches an infinite multitude of $x^{(l)}$, all of them mapped into a single CIE point. All these different versions of $x^{(l)}$, however, are equivalent: they all match precisely the same colors on the right. Another example: a stimulus of luminance level L_1 and size s_1 can have the same (subjective) brightness as a stimulus of some other luminance L_2 and size s_2, regardless of whether the two stimuli belong to the 'left' or 'right' stimulus area. One would expect then that $(L_2, s_2)^{(l)}\,\mathrm{M}\,(L, s)^{(r)}$ if and only if $(L_1, s_1)^{(l)}\,\mathrm{M}\,(L, s)^{(r)}$. That is, a given right-hand stimulus would match more than one left-hand stimulus. It would be reasonable to expect, however, that then $(L_1, s_1)^{(l)}$ and $(L_2, s_2)^{(l)}$ are equivalent, *i.e.*, that it is impossible for one of them to match and the other to not match one and the same stimulus on the right. This is precisely what RM2 posits: the uniqueness of matches *up to equivalence*.

The following proposition contains two most important consequences of RM1-RM2.

Proposition 1. *Assuming* M *satisfies* RM1 *and* RM2, *we have*

(1) $a^{(l)}\,\mathrm{M}\,b^{(r)} \Longleftrightarrow b^{(r)}\,\mathrm{M}\,a^{(l)}$,

(2a) $a^{(l)}\,\mathrm{M}\,b^{(r)} \wedge b^{(r)}\,\mathrm{M}\,c^{(l)} \wedge c^{(l)}\,\mathrm{M}\,d^{(r)} \Longrightarrow a^{(l)}\,\mathrm{M}\,d^{(r)}$,
(2b) $a^{(r)}\,\mathrm{M}\,b^{(l)} \wedge b^{(l)}\,\mathrm{M}\,c^{(r)} \wedge c^{(r)}\,\mathrm{M}\,d^{(l)} \Longrightarrow a^{(r)}\,\mathrm{M}\,d^{(l)}$
for all $a^{(l)}, b^{(r)}, c^{(r)}, d^{(l)}$.

Proof. To prove (1), suppose $a^{(l)}\,\mathrm{M}\,b^{(r)}$. By RM1, there exists some $e^{(l)}$ with $e^{(l)}\,\mathrm{M}\,b^{(r)}$ and $b^{(r)}\,\mathrm{M}\,e^{(l)}$. By RM2, the first of these relations implies that $e^{(l)}\,\mathrm{E}\,a^{(l)}$, and by definition of E) the second implies that $b^{(r)}\,\mathrm{M}\,a^{(l)}$. Thus $a^{(l)}\,\mathrm{M}\,b^{(r)} \Longrightarrow b^{(r)}\,\mathrm{M}\,a^{(l)}$; the reverse implication is proved symmetrically.

We now prove (2a), the proof of (2b) being symmetric. Assume that $a^{(l)}\,\mathrm{M}\,b^{(r)}$, $b^{(r)}\,\mathrm{M}\,c^{(l)}$, and $c^{(l)}\,\mathrm{M}\,d^{(r)}$. By symmetry of M we have $d^{(r)}\,\mathrm{M}\,c^{(l)}$, and since $b^{(r)}\,\mathrm{M}\,c^{(l)}$ RM2 implies $b^{(r)}\,\mathrm{E}\,d^{(r)}$. Since $a^{(l)}\,\mathrm{M}\,b^{(r)}$, $a^{(l)}\,\mathrm{M}\,d^{(r)}$ by definition of E. □

The proposition says that if one accepts RM1 and RM2 (which are in agreement with, or at least do not contradict, what we know about human comparative judgments), then the (idealized) 'matching' relation designed to generalize psychophysical matching ought to be *symmetric* in the sense of (1) and satisfy the notion of *tetradic transitivity*, (2). Let us briefly comment on each of these.

Regarding symmetry, it is very important to note that the values x and y in the expression $x^{(l)}\,\mathrm{M}\,y^{(r)} \Longleftrightarrow y^{(r)}\,\mathrm{M}\,x^{(l)}$ remain in their respective stimulus areas on both sides. The symmetry condition does not necessarily allow for the exchange of values between the two stimulus areas,

$$x^{(l)}\,\mathrm{M}\,y^{(r)} \Longleftrightarrow y^{(l)}\,\mathrm{M}\,x^{(r)}.$$

The naive notion of symmetry represented by this statement is definitely not a general rule.[11] The symmetry in the sense of (1), on the other hand, is supported by all available empirical evidence (Dzhafarov, 2002; Dzhafarov & Colonius, 2006) and underlies the very language of psychophysical research dealing with matching-type relations. A psychophysicist is likely to consider the description

[11]Thus, in one of the same-different discrimination experiments described in Dzhafarov & Colonius (2005), a right-hand segment of length x happens to match a left-hand segment of length $x-2$, if measured in minutes of arc. So, a $17^{(r)}$ min arc segment and a $15^{(l)}$ min arc one match each other. But, clearly, $15^{(r)}$ min arc and $17^{(l)}$ min arc do not match: rather the former of the two is matched by $13^{(l)}$ min arc. One can list a host of such illustrations involving what in psychophysics is called *constant error*.

> the observer adjusted the right-hand stimulus until its appearance matched that of the fixed stimulus on the left

as saying precisely the same as

> the observer adjusted the right-hand stimulus until the fixed stimulus on the left matched its appearance

and precisely the same as

> the observer adjusted the right-hand stimulus until its appearance and that of the fixed stimulus on the left matched each other.

Regarding now the tetradic form of transitivity in Proposition 1, it is easily seen that the 'ordinary', triadic transitivity is simply false (or unformulable) if one deals with two stimulus areas: if $a^{(l)}\,\mathrm{M}\,b^{(r)}$ and $b^{(r)}\,\mathrm{M}\,c^{(l)}$, it is never true that $a^{(l)}\,\mathrm{M}\,c^{(l)}$ since only stimuli from different stimulus areas can be compared. Thus M can be said to be transitive in the tetradic sense but *antitransitive* in the triadic sense. It is easy to see, and will be rigorously demonstrated in the next subsection, that the tetradic transitivity is all one needs to rule out the existence of comparative soritical sequences in the case involving just two distinct stimulus areas (not necessarily the 'left' and 'right' used here for concreteness only).

If the number of stimulus areas is greater than two, the analysis above does apply, of course, to any two of them. In addition, however, for any three distinct stimulus areas (let us denote them 1, 2, and 3) one can formulate the familiar triadic transitivity property,

$$a^{(1)}\,\mathrm{M}\,b^{(2)} \wedge b^{(2)}\,\mathrm{M}\,c^{(3)} \Longrightarrow a^{(1)}\,\mathrm{M}\,dc^{(3r)}.$$

This property can be derived from appropriate reformulations of RM1 and RM2 for three distinct stimulus areas. We forgo this task, however, as it is subsumed by the formal treatment presented next, which applies to an arbitrary set of stimulus areas.

Formal theory of regular well-matched stimulus spaces

We work throughout with a set $S \times \Omega$ in which at least two stimulus values are paired with at least two stimulus areas. We endow $S \times \Omega$ with

(1) a binary relation M such that $\neg a^{(\omega)}\,\mathrm{M}\,b^{(\omega)}$ for all $a^{(\omega)}, b^{(\omega)} \in S \times \Omega$;
(2) a binary relation E such that $a^{(\omega)}\,\mathrm{E}\,b^{(\omega')}$ if and only if

$$c^{(\iota)}\,\mathrm{M}\,a^{(\omega)} \Longleftrightarrow c^{(\iota)}\,\mathrm{M}\,b^{(\omega')}$$

for all $a^{(\omega)}, b^{(\omega')}, c^{(\iota)} \in S \times \Omega$.

Since E is uniquely defined in terms of M, we refer to the space $(S \times \Omega, \mathrm{M}, \mathrm{E})$ by the more economical $(S \times \Omega, \mathrm{M})$.[12] We omit the simple proof that E is an equivalence relation on $S \times \Omega$.

Definition 5. A sequence $x_1^{(\omega_1)}, \ldots, x_n^{(\omega_n)}$ in a space $(S \times \Omega, \mathrm{M})$ is called
(1) *chain-matched* if $x_i^{(\omega_i)}\,\mathrm{M}\,x_{i+1}^{(\omega_{i+1})}$ for $i = 1, \ldots, n-1$;
(2) *well-matched* if $\omega_i \neq \omega_j \Longrightarrow x_i^{(\omega_i)}\,\mathrm{M}\,x_j^{(\omega_j)}$ for all $i, j \in \{1, \ldots n\}$;
(3) *soritical* if it is chain-matched, $\omega_1 \neq \omega_n$ and $\neg x_1^{(\omega_1)}\,\mathrm{M}\,x_n^{(\omega_n)}$.

In (3) we easily recognize a formal version of what we earlier defined as comparative soritical sequences. Soritical sequences are clearly always chain-matched and never well-matched.[13] It is also clear that there are no soritical sequences with just two elements, and that all soritical sequences consisting of three elements are of the form $a^{(\alpha)}, b^{(\beta)}, c^{(\gamma)}$ with $\{\alpha, \beta, \gamma\}$ pairwise distinct. Longer soritical sequences, as it turns out, can always be reduced to one of two types (illustrated in Fig. 1.2): three-element sequences like the one just mentioned, and four-element sequences with two alternating stimulus areas, $a^{(\alpha)}$, $b^{(\beta)}$, $c^{(\alpha)}$, $d^{(\beta)}$.

Lemma 4. *If $x_1^{(\omega_1)}, \ldots, x_n^{(\omega_n)}$ in a space $(S \times \Omega, \mathrm{M})$ is a soritical sequence, then it contains either a triadic soritical subsequence $a^{(\alpha)}, b^{(\beta)}, c^{(\gamma)}$ or a tetradic soritical subsequence $a^{(\alpha)}, b^{(\beta)}, c^{(\alpha)}, d^{(\beta)}$.*

Proof. Let $x_{i_1}^{(\omega_{i_1})}, \ldots, x_{i_m}^{(\omega_{i_m})}$ be a soritical subsequence of our sequence having the shortest possible length. If there exists an ℓ such that $1 <$

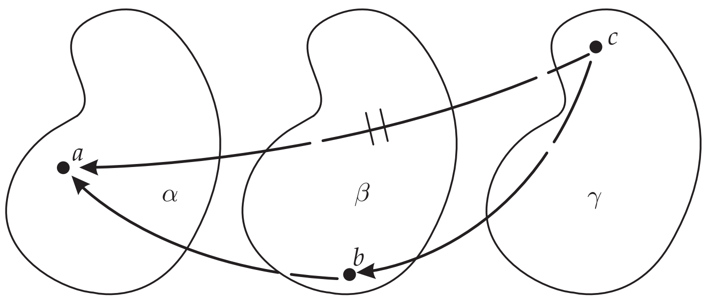
[12] In dealing with both E and M, and treating them as two interrelated but different relations, we follow Goodman (1951/1997).

[13] This shows that chain-matchedness does not imply well-matchedness. That the reverse implication also does not hold can be seen by considering any sequence of the form $x_1^{(\omega)}, \ldots, x_n^{(\omega)}$. Thus, chain-matchedness and well-matchedness are logically independent conditions.

$\ell < m$ and $\omega_{i_1} \neq \omega_{i_\ell} \neq \omega_{i_m}$ then it must be that $x_{i_1}^{(\omega_{i_1})} \mathrm{M}\, x_{i_\ell}^{(\omega_{i_\ell})}$: otherwise $x_{i_1}^{(\omega_{i_1})}, \ldots, x_{i_\ell}^{(\omega_{i_\ell})}$ would be yet a shorter soritical subsequence of the original sequence. Similarly, it must be that $x_{i_\ell}^{(\omega_{i_\ell})} \mathrm{M}\, x_{i_m}^{(\omega_{i_m})}$. Hence,

$$(a^{(\alpha)}, b^{(\beta)}, c^{(\gamma)}) = (x_{i_1}^{(\omega_{i_1})}, x_{i_\ell}^{(\omega_{i_\ell})}, x_{i_m}^{(\omega_{i_m})})$$

is a triadic subsequence of the kind desired. If no such ℓ exists, then



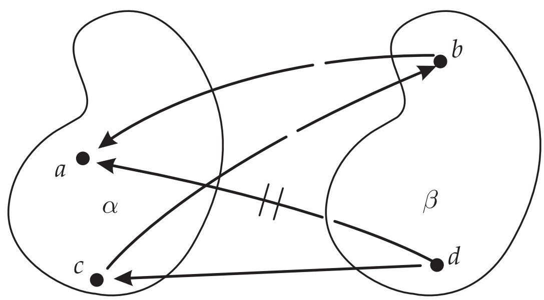


Figure 1.2: Examples of a triadic soritical sequence (top) and a tetradic soritical sequence (bottom). According to Lemma 4, at least one of these can be found as a subsequence in any soritical sequence. The outlined areas represent stimulus areas: α, β, γ in the top illustration and α, β in the bottom one. An arrow (resp., interrupted arrow) drawn from one point to another indicates that the latter point is matched (resp., not matched) by the former. Thus, in the top illustration, $b^{(\beta)}$ matches $a^{(\alpha)}$, $c^{(\gamma)}$ matches $b^{(\beta)}$, but $c^{(\gamma)}$ does not match $a^{(\alpha)}$. In the bottom illustration, $b^{(\beta)}$ matches $a^{(\alpha)}$, $c^{(\alpha)}$ matches $b^{(\beta)}$, $d^{(\beta)}$ matches $c^{(\alpha)}$, but $d^{(\beta)}$ does not match $a^{(\alpha)}$. According to Theorem 3 (p. 26), neither of these scenarios (and hence no soritical sequence) is possible in a regular well-matched space (Definition 6).

it must be that $m \geq 4$ and that $\omega_{i_1} = \omega_{i_3} \neq \omega_{i_2} = \omega_{i_m}$. Again, the choice of $x_{i_1}^{(\omega_{i_1})}, \ldots, x_{i_m}^{(\omega_{i_m})}$ as the shortest soritical subsequence ensures that $x_{i_3}^{(\omega_{i_3})} \mathrm{M}\, x_{i_m}^{(\omega_{i_m})}$, so in this case

$$(a^{(\alpha)}, b^{(\beta)}, c^{(\alpha)}, d^{(\beta)}) = (x_{i_1}^{(\omega_{i_1})}, x_{i_2}^{(\omega_{i_2})}, x_{i_3}^{(\omega_{i_3})}, x_{i_m}^{(\omega_{i_m})})$$

is a tetradic soritical subsequence of our sequence. □

Our goal being to dissolve the comparative sorites 'paradox' in systems resembling human comparative judgments, we would like to restrict our interest from arbitrary spaces to those in which we can formalize (suitably generalized versions of) the conditions RM1 and RM2 discussed in the previous subsection. To this end, we introduce the following two types of spaces.

Definition 6. We call $(S \times \Omega, \mathrm{M})$ a

(1) *well-matched space* if, for any sequence $\alpha, \beta, \gamma \in \Omega$ and any $a^{(\alpha)} \in S \times \Omega$, there is a well-matched sequence $a^{(\alpha)}, b^{(\beta)}, c^{(\gamma)}$;

(2) *regular space* if, for any $a^{(\alpha)}, b^{(\alpha)}, c^{(\beta)} \in S \times \Omega$ with $\alpha \neq \beta$,

$$a^{(\alpha)} \mathrm{M}\, c^{(\beta)} \wedge b^{(\alpha)} \mathrm{M}\, c^{(\beta)} \Longrightarrow a^{(\alpha)} \mathrm{E}\, b^{(\alpha)}.$$

In the concept of regularity, we clearly see the formalization of RM2. The following lemma shows that, similarly, the notion of well-matchedness suffices for the formalization of RM1. It also shows that in such spaces the relation E behaves as we expect it to.

Lemma 5. *Let $(S \times \Omega, \mathrm{M})$ be a well-matched space. (1) For any $\alpha, \beta \in \Omega$ and any $a^{(\alpha)} \in S \times \Omega$, there exists $b^{(\beta)} \in S \times \Omega$ such that $a^{(\alpha)} \mathrm{M}\, b^{(\beta)}$ and $b^{(\beta)} \mathrm{M}\, a^{(\alpha)}$. (2) For any $a^{(\alpha)}, b^{(\beta)} \in S \times \Omega$, if $a^{(\alpha)} \mathrm{E}\, b^{(\beta)}$ holds then $\alpha = \beta$.*

Proof. For (1), consider a sequence α, β, β and apply Definition 6(1) (notice that no assumption is made in that definition that α, β, γ need to be pairwise distinct). To prove (2), suppose for a contradiction that $a^{(\alpha)} \mathrm{E}\, b^{(\beta)}$ and $\alpha \neq \beta$. Then by part (1), we can find some $c^{(\alpha)}$ with $c^{(\alpha)} \mathrm{M}\, b^{(\beta)}$ and $b^{(\beta)} \mathrm{M}\, c^{(\alpha)}$. By definition of E, this implies that $c^{(\alpha)} \mathrm{M}\, a^{(\alpha)}$, which contradicts the definition of M. □

It follows that if comparative judgments satifying RM1 and RM2 are to serve as the model for our formal analysis, then we should apply our analysis to spaces which are both well-matched and regular.[14] The following lemma provides us with additional information about the structure of such spaces. Part (1) is the analog of Proposition 1(1); parts (2) and (3) are obvious and are stated mostly for convenience of reference in the next subsection.

Lemma 6. *Let* $(S \times \Omega, \mathrm{M})$ *be a regular well-matched space, and let* $a^{(\alpha)}$, $b^{(\beta)}, c^{(\alpha)}, d^{(\beta)} \in S \times \Omega$. *Then*
(1) $a^{(\alpha)}\,\mathrm{M}\,b^{(\beta)} \iff b^{(\beta)}\,\mathrm{M}\,a^{(\alpha)}$;
(2) *if either* $a^{(\alpha)}\,\mathrm{M}\,b^{(\beta)} \wedge c^{(\alpha)}\,\mathrm{M}\,b^{(\beta)}$ *or* $b^{(\beta)}\,\mathrm{M}\,a^{(\alpha)} \wedge b^{(\beta)}\,\mathrm{M}\,c^{(\alpha)}$, *then* $a^{(\alpha)}\,\mathrm{E}\,c^{(\alpha)}$;
(3) *if* $a^{(\alpha)}\,\mathrm{E}\,c^{(\alpha)} \wedge b^{(\beta)}\,\mathrm{E}\,d^{(\beta)}$ *then* $a^{(\alpha)}\,\mathrm{M}\,b^{(\beta)} \iff c^{(\alpha)}\,\mathrm{M}\,d^{(\beta)}$.

Proof. Part (1) is proved exactly as Proposition 1(1), only replacing references to RM1 by those to Lemma 5(1) and replacing references to RM2 by those to the regularity of the space. Parts (2) and (3) are straightforward consequences of Definition 2 and part (1). □

Dissolving the comparative sorites 'paradox'

The main result of the section is in the theorem presented next. As mentioned above, soritical sequences are always chain-matched and never well-matched, so the theorem in particular implies that soritical sequences do not exist in regular well-matched spaces.

Theorem 3. *Let* $(S \times \Omega, \mathrm{M})$ *be a regular well-matched space. Then any chain-matched sequence in this space is well-matched.*

Proof. In view of Lemma 4, it suffices to prove the result for all soritical sequences of the form $a^{(\alpha)}, b^{(\beta)}, c^{(\gamma)}$ and $a^{(\alpha)}, b^{(\beta)}, c^{(\alpha)}, d^{(\beta)}$. We begin with the former.

Suppose $a^{(\alpha)}, b^{(\beta)}, c^{(\gamma)}$ is a chain-matched sequence. This means $a^{(\alpha)}\,\mathrm{M}\,b^{(\beta)} \wedge b^{(\beta)}\,\mathrm{M}\,c^{(\gamma)}$. All we have to prove is that then $a^{(\alpha)}\,\mathrm{M}\,c^{(\gamma)}$, as the rest of the matches in $a^{(\alpha)}, b^{(\beta)}, c^{(\gamma)}$ then obtain by symmetry of M. By Definition 6, there exists a well-matched sequence $x^{(\alpha)}, b^{(\beta)}, y^{(\gamma)}$. Since $a^{(\alpha)}\,\mathrm{M}\,b^{(\beta)} \wedge x^{(\alpha)}\,\mathrm{M}\,b^{(\beta)}$, it follows by Lemma 6(2) that $a^{(\alpha)}\,\mathrm{E}\,x^{(\alpha)}$. Since $b^{(\beta)}\,\mathrm{M}\,c^{(\gamma)} \wedge b^{(\beta)}\,\mathrm{M}\,y^{(\gamma)}$, it follows similarly that $c^{(3)}\,\mathrm{E}\,y^{(3)}$. Since $x^{(\alpha)}, b^{(\beta)}, y^{(\gamma)}$ is well-matched, we have $x^{(\alpha)}\,\mathrm{M}\,y^{(\gamma)}$, and so, by Lemma 6(3), $a^{(1)}\,\mathrm{M}\,c^{(3)}$.

[14]These are independent assumptions. It is easy to construct toy examples demonstrating that a well-matched space need not be regular, and vice-versa.

Now suppose $a^{(\alpha)}, b^{(\beta)}, c^{(\alpha)}, d^{(\beta)}$ is a chain-matched sequence, so $a^{(\alpha)}\,\mathrm{M}\,b^{(\beta)} \wedge b^{(\beta)}\,\mathrm{M}\,c^{(\alpha)} \wedge c^{(\alpha)}\,\mathrm{M}\,d^{(\beta)}$. We have to show that $a^{(\alpha)}\,\mathrm{M}\,d^{(\beta)}$. By Definition 6, there exists a well-matched sequence $x^{(\alpha)}$, $b^{(\beta)}$, $y^{(\alpha)}$, $z^{(\beta)}$. Since $b^{(\beta)}\,\mathrm{M}\,c^{(\alpha)} \wedge b^{(\beta)}\,\mathrm{M}\,y^{(\alpha)}$, it follows by Lemma 6(2) that $c^{(\alpha)}\,\mathrm{E}\,y^{(\alpha)}$. Then we should have $y^{(\alpha)}\,\mathrm{M}\,d^{(\beta)}$ (by Lemma 6(3), because $c^{(\alpha)}\,\mathrm{M}\,d^{(\beta)}$). Since $y^{(\alpha)}\,\mathrm{M}\,d^{(\beta)} \wedge y^{(\alpha)}\,\mathrm{M}\,z^{(\beta)}$, we have by Lemma 6(2) that $d^{(\beta)}\,\mathrm{E}\,z^{(\beta)}$. By the same lemma, we also have $a^{(\alpha)}\,\mathrm{E}\,x^{(\alpha)}$, because $a^{(\alpha)}\,\mathrm{M}\,b^{(\beta)} \wedge x^{(\alpha)}\,\mathrm{M}\,b^{(\beta)}$. But now $a^{(\alpha)}\,\mathrm{E}\,x^{(\alpha)}$ and $d^{(\beta)}\,\mathrm{E}\,z^{(\beta)}$, so from the fact that $x^{(\alpha)}\,\mathrm{M}\,z^{(\beta)}$ it follows that $a^{(\alpha)}\,\mathrm{M}\,d^{(\beta)}$, by Lemma 6(3). $\square$

Corollary 2. *Any chain-matched sequence in a regular well-matched space is well-matched: one cannot form a soritical sequence in such a space.*

We conclude with a method of merging the antireflexive relation M and the equivalence relation E into a single identity relation EM. To this end, we define the notion of 'canonical labeling'. The idea is very simple: given a regular well-matched space $S \times \Omega$, any two equivalent stimuli $a^{(\omega)}$ and $b^{(\omega)}$ in any stimulus area ω can be assigned one and the same *label* (say, $\mathbf{x}$). Then every new label in any one stimulus area will match and be matched by one and only one label in any other stimulus area—and then it is possible to assign the same label $\mathbf{x}$ to all stimuli in all stimulus areas which match (and are matched by) $\mathbf{x}^{(\omega)}$. The resulting simplicity is the reward: for any two stimulus areas ω and ω', any 'relabeled stimulus' $\mathbf{x}^{(\omega)}$ matches the 'relabeled stimulus' $\mathbf{x}^{(\omega')}$ and none other; and in any given stimulus area any $\mathbf{x}^{(\omega)}$ is only equivalent to itself. If one now 'merges' the relations E and M into a single relation EM, the latter is simply the indicator of the equality of labels and is therefore reflexive, symmetric, and transitive:

$$\mathbf{a}^{(\alpha)}\mathrm{EM}\mathbf{b}^{(\beta)} \iff \mathbf{a} = \mathbf{b},$$

where α and β need not be distinct. The formal procedure described below effects the canonical (re)labeling by means of a single function cal (from 'canonical labeling') applied to all stimuli in the space.

Canonical representation is not a return to the naive idea that every stimulus matches 'itself'. One cannot dispense with the notion of a stimulus area: for distinct ω and ω' and one and the same label $\mathbf{x}$, the original identities (*e.g.*, conventional physical descriptions) of $\mathbf{x}^{(\omega)}$ and $\mathbf{x}^{(\omega')}$ are generally different. In fact, $\mathbf{x}^{(\omega)}$ and $\mathbf{x}^{(\omega')}$ designate two equivalence classes of stimuli, whose values may be non-overlapping (partially or completely).

Definition 7. A surjective function

$$\mathrm{cal}:\ S\times\Omega\to\mathbf{S}$$

is called a *canonical labeling* of a regular well-matched space $(S\times\Omega,\ \mathrm{M}\,)$ (and $\mathbf{S}$ is called a *set of canonical labels*) if, for any $a^{(\alpha)},b^{(\beta)}\in S\times\Omega$,

$$\mathrm{cal}(a^{(\alpha)})=\mathrm{cal}(b^{(\beta)})\Longleftrightarrow\begin{cases}a^{(\alpha)}\,\mathrm{E}\,b^{(\beta)} & \text{if }\alpha=\beta\\ a^{(\alpha)}\,\mathrm{M}\,b^{(\beta)} & \text{if }\alpha\neq\beta\end{cases}.$$

Theorem 4. *A canonical labeling function* cal *exists for any regular well-matched space* $(S\times\Omega,\ \mathrm{M}\,)$*. If* $\mathrm{cal}:\ S\times\Omega\to\mathbf{S}$ *and* $\mathrm{cal}^*:\ S\times\Omega\to\mathbf{S}^*$ *are such functions, then* $\mathrm{cal}^*\equiv h\circ\mathrm{cal}$*, where h is a bijection* $\mathbf{S}\to\mathbf{S}^*$*.*

Proof. For each $x^{(\omega)}\in S\times\Omega$, let

$$N_{x^{(\omega)}}=\{y^{(\iota)}\in S\times\Omega\colon x^{(\omega)}\,\mathrm{M}\,y^{(\iota)}\vee x^{(\omega)}\,\mathrm{E}\,y^{(\iota)}\}.$$

Since $(S\times\Omega,M)$ is regular and well-matched, it is easy to see that if $N_{x^{(\omega)}}\cap N_{y^{(\iota)}}\neq\varnothing$ for some $x^{(\omega)},y^{(\iota)}\in S\times\Omega$, then $N_{x^{(\omega)}}=N_{y^{(\iota)}}$. Setting $\mathbf{S}=\{N_{x^{(\omega)}}:x^{(\omega)}\in S\times\Omega\}$, we define $\mathrm{cal}:\ S\times\Omega\to\mathbf{S}$ by $\mathrm{cal}(x^{(\omega)})=N_{x^{(\omega)}}$, and this function clearly satisfies Definition 7. If now $\mathrm{cal}^*:\ S\times\Omega\to\mathbf{S}^*$ is another canonical labeling function for $(S\times\Omega,M)$, define $h:\ \mathbf{S}\to\mathbf{S}^*$ by $h(N_{x^{(\omega)}})=\mathrm{cal}^*(x^{(\omega)})$. Since $\mathrm{cal}^*(x^{(\omega)})=\mathrm{cal}^*(y^{(\iota)})$ if and only if $x^{(\omega)}\,\mathrm{M}\,y^{(\iota)}$ or $x^{(\omega)}\,\mathrm{E}\,y^{(\iota)}$, it follows that h is well-defined and injective, while its surjectivity follows immediately from that of cal^*. Clearly, $\mathrm{cal}^*=h\circ\mathrm{cal}$, whence the proof is complete. □

Consider now the set $\mathbf{S}\times\Omega$. As before, let us use the notation $\mathbf{x}^{(\omega)}$ for $(\mathbf{x},\omega)$.

Definition 8. Given a regular well-matched space $(S\times\Omega,\ \mathrm{M}\,)$ and a canonical labeling function $\mathrm{cal}:\ S\times\Omega\to\mathbf{S}$, for any $\mathbf{a},\mathbf{b}\in\mathbf{S}$ and $\alpha,\beta\in\Omega$, we say that $\mathbf{b}^{(\beta)}$ *matches* $\mathbf{a}^{(\alpha)}$, and write $\mathbf{a}^{(\alpha)}\mathrm{EM}\mathbf{b}^{(\beta)}$, if

$$a^{(\alpha)}\,\mathrm{M}\,b^{(\beta)}\vee a^{(\alpha)}\,\mathrm{E}\,b^{(\beta)}$$

for some $a^{(\alpha)}$and $b^{(\beta)}$ in $S\times\Omega$ such that $\mathrm{cal}(a^{(\alpha)})=\mathbf{a}$ and $\mathrm{cal}(b^{(\beta)})=\mathbf{b}$. We refer to $(\mathbf{S}\times\Omega,\mathrm{EM})$ as a *canonical comparison space*.

Theorem 5. *Given any regular well-matched space* $(S \times \Omega, \mathrm{M})$ *and any canonical labeling function* $\mathrm{cal} : S \times \Omega \to \mathbf{S}$, *for any* $\mathbf{a}, \mathbf{b} \in \mathbf{S}$ *and* $\alpha, \beta \in \Omega$ *we have*

$$\mathbf{a}^{(\alpha)} \mathrm{EM} \mathbf{b}^{(\beta)} \Longleftrightarrow \mathbf{a} = \mathbf{b}.$$

Hence EM *is an equivalence relation on* $\mathbf{S} \times \Omega$.

Proof. By Definition 8, $\mathbf{a}^{(\alpha)} \mathrm{EM} \mathbf{b}^{(\beta)}$ means that for some $a^{(\alpha)}$ with $\mathrm{cal}(a^{(\alpha)}) = \mathbf{a}$ and $b^{(\beta)}$ with $\mathrm{cal}(b^{(\beta)}) = \mathbf{b}$, either $a^{(\alpha)} \,\mathrm{M}\, b^{(\beta)}$ (which implies $\alpha \neq \beta$) or $a^{(\alpha)} \,\mathrm{E}\, b^{(\beta)}$ (implying $\alpha = \beta$). But by Definition 7,

$$\left.\begin{array}{ll} a^{(\alpha)} \,\mathrm{E}\, b^{(\beta)} & (\alpha = \beta) \\ a^{(\alpha)} \,\mathrm{M}\, b^{(\beta)} & (\alpha \neq \beta) \end{array}\right\} \Longleftrightarrow \mathrm{cal}(a^{(\alpha)}) = \mathrm{cal}(b^{(\beta)}).$$

This proves the theorem. □

Corollary 3. *With* EM *in place of* M *and* $\mathbf{S}$ *in place of* S, *any canonical comparison space* $(\mathbf{S} \times \Omega, \mathrm{EM})$ *is a regular well-matched space.*

Conclusion

We have approached soritical arguments and (allegedly) soritical phenomena within a broadly understood behavioral framework, in terms of stimuli acting upon a system (such as a biological organism, a group of people, or a set of normative linguistic rules) and the responses they evoke.

The classificatory sorites, dating back to Eubulides of the Megarian school, is about the identity of or difference between the effects of stimuli which differ 'only microscopically'. We have formulated the notions and assumptions underlying this variety of sorites in a highly general mathematical language, and we have shown that the 'paradox' is dissolved on grounds unrelated to vague predicates or other linguistic issues traditionally associated with it. If stimulus effects are properly defined (*i.e.*, if they are uniquely determined by stimuli), and if the space of the stimuli is endowed with appropriate closeness and connectedness properties, then this space must contain points in every vicinity of which, 'however small', the stimulus effect is not constant. This conclusion clashes with the common but nonetheless false intuition that a 'macroscopic' system cannot 'react differently' to two 'microscopically different' stimuli. In fact, a non-constant stimulus effect upon a system can only be insensitive to small differences in stimuli if the closeness structure which is used to define very close stimuli does not render the space of stimuli appropriately connected, and in this case we have no 'paradox'. The 'paradox' cannot even be formulated using response properties that are not true stimulus effects, *i.e.*, are not uniquely determined by stimuli. This is the reason the classificatory sorites is not related to the issue of vagueness in human responses to stimuli: 'vague predicates' are always assigned inconsistently, whatever other properties they may be thought to have.

The comparative sorites (also known in the literature as 'observational') is very different from the classificatory one. Here, it has been discussed in terms of a system mapping pairs of stimuli into a binary response characteristic whose values are uniquely determined by stimulus pairs and are interpretable as the complementary relations 'match' and 'do not match' (overall or in some respect). The comparative sorites is about hypothetical sequences of stimuli in which every two successive elements are mapped into the relation 'match', while the pair comprised of the first and the last elements of the sequence is mapped into 'do not match'. Although soritical sequences of this kind are logically possible, we have argued that insofar as human comparative judgments are concerned, their existence is far from being a well-known, let alone obvious, empirical fact. Rather it is a naive theoretical idea that overlooks the fundamental notion of stimulus areas and the necessity of defining the matching relation so that it is uniquely determined by stimulus pairs (which is critical in view of the probabilistic nature of comparative judgments in humans). Moreover, the comparative soritical sequences are excluded by the principle of Regular Mediality/Minimality proposed for human comparative judgments in a context unrelated to soriticial issues. In this chapter we have generalized this principle into the mathematical notion of regular well-matched spaces of stimulus values paired with stimulus areas. The matching relation in such spaces is irreflexive, symmetric, and transitive in the tetradic or triadic sense, depending as we deal with two or more than two stimulus areas, respectively.

Acknowledgments

The first author's work was supported by NSF grant SES 0620446 and AFOSR grants FA9550-06-1-0288 and FA9550-09-1-0252. The second author's work was partially supported by an NSF Graduate Research Fellowship.

1.1 Bibliography

Armstrong, D. M. (1968). *A Materialist Theory of the Mind*. London: Routledge and Kegan Paul.

Dummett, M. (1975). Wang's paradox. *Synthese*, *30*, 301–324.

Dzhafarov, E. N. (2002). Multidimensional Fechnerian scaling: Pairwise comparisons, regular minimality, and nonconstant self-similarity. *Journal of Mathematical Psychology*, *46*, 583–608.

Dzhafarov, E. N. (2003). Thurstonian-type representations for 'same-different' discriminations: Deterministic decisions and independent images. *Journal of Mathematical Psychology*, *47*, 208–228.

Dzhafarov, E. N., & Colonius, H. (2005). Psychophysics without physics: A purely psychological theory of Fechnerian scaling in continuous stimulus spaces. *Journal of Mathematical Psychology*, 1–50.

Dzhafarov, E. N., & Colonius, H. (2006). Regular Minimality: A fundamental law of discrimination. In H. Colonius & E. Dzhafarov (Eds.), *Measurement and Representation of Sensations* (pp. 1–46). Mahwah: Lawrence Erlbaum Associates, Inc.

Dzhafarov, E. N., & Dzhafarov, D. D. (2010a). Sorites without vagueness I: Classificatory sorites. *Theoria*, *76*(1), 4–24.

Dzhafarov, E. N., & Dzhafarov, D. D. (2010b). Sorites without vagueness II: Comparative sorites. *Theoria*, *76*(1), 25–53.

Fréchet, M. (1918). Sur la Notion de Voisinage Dans Les Ensembles Abstrait [On the notion of neighborhood in abstract sets]. *Bulletin des Sciences Mathématiques*, *42*, 138–156.

Goodman, N. (1951/1997). *The Structure of Appearance*. Dordrecht: Reidel.

Graff, D. (2001). Phenomenal continua and the sorites. *Mind*, *110*, 905–935.

Hardin, C. L. (1988). Phenomenal colors and sorites. *Noûs*, *22*, 213–234.

Jackson, F., & Pinkerton, R. J. (1973). On an argument against sensory items. *Mind*, *82*, 269–272.

Kujala, J., & Dzhafarov, E. N. (2008). On minima of discrimination functions. *Journal of Mathematical Psychology*, *52*, 116–127.

Luce, R. D. (1956). Semiorders and a theory of utility discrimination. *Econometrica*, *24*, 178–191.

Peirce, C. S. (1901/1960). Vague. In J. M. Baldwin (Ed.), *Dictionary of philosophy and psychology* (p. 748). New York: MacMillan.

Sorensen, R. A. (1988a). *Blindspots*. Oxford: Clarendon Press.

Sorensen, R. A. (1988b). Vagueness, blurriness, and measurement. *Synthese*, *75*, 45–82.

Varzi, A. C. (2003). Cut-offs and their neighbors. In J. Beall (Ed.), *Liars and heaps* (pp. 24–38). Oxford: Oxford University Press.

Williamson, T. (1994). *Vagueness*. London: Routledge.

Williamson, T. (2000). *Knowledge and its limits*. Oxford: Oxford University Press.

Wright, C. (1975). On the coherence of vague predicates. *Synthese*, *30*, 325–365.

Yonge, C. D. (1901). *The Lives and Opinions of Eminent Philosophers, by Diogenes Laërtius*. London: G. Bell & Sons.