To appear in S. Oms & E. Zardini eds., The Sorites Paradox
Cambridge University Press, 2018
(This is only a preview, pp. 1-2/28, please contact first author for full version)

THE SORITES PARADOX IN PSYCHOLOGY

PAUL ÉGRÉ, DAVID RIPLEY, STEVEN VERHEYEN

Abstract. This chapter examines some aspects of the influence of the sorites paradox in psychology. Section 1 starts out with a brief discussion of the analysis of slippery slope arguments in the psychology of reasoning, to introduce the relevance of probabilistic considerations in that domain. We then devote most of this chapter to the analysis in psychophysics and in the psychology of concepts of the complex relationship between discrimination and categorization for items that differ very little. Section 2 emphasizes the centrality of probabilistic modeling to represent the way in which small differences between stimuli affect decisions of membership under a common category. Section 3 focuses on experimental data concerning unordered transitions between prototypes, then section 4 looks at data concerning ordered transitions between prototypes (dynamic sorites).

This chapter examines some areas of theoretical and experimental psychology in which the sorites paradox has had an influence or has been an object of study. Our goal is to show not only different manifestations of the sorites in psychology, but also how psychological modeling and behavioral data can cast light on the puzzle raised by the paradox.

The first aspect we consider concerns the *psychology of reasoning and argumentation*. Sorites arguments are often conflated with 'slippery slope arguments', typically used *a contrario* to argue that a line should be drawn at a specific location of a vague domain, on pain of reaching an undesirable or absurd outcome, or alternatively that no line can be drawn at all. In Section 1 we start out with a brief history and overview of work done on slippery slope arguments, to highlight that such arguments are not intrinsically wrong: fundamentally they are inductive arguments, whose acceptability depends on the strength of the relation between the antecedent and the consequent of their conditional premises, and on the utility attached to specific outcomes. As such, slippery slope arguments tend to be handled in a probabilistic framework.

The second and more significant area of influence we consider, concerns the study of similarity in *psychophysics* and in the *psychology of concepts*. The main premise of a sorites argument involves the notion of sufficient similarity between objects, and states that if two objects are sufficiently similar, they must produce identical judgments as to whether some property applies or not. Section 2 presents some influential accounts of the relation between discrimination and categorization in psychophysics, and underscores the centrality of probabilistic modeling to deal with sorites-susceptible predicates quite generally.

We distinguish, following Raffman (1994) and Dzhafarov and Dzhafarov (2012), two versions of the main premise of the sorites, one pertaining to discrimination (same vs. different comparison task), and one pertaining to categorization (assignment under a common

lexical category). We look first at the psychology of discrimination, and at how the notion of just noticeable difference introduced by Fechner can be related to the tolerance principle, namely the idea that some differences can be so small as to make no difference in terms of discrimination. We then look at the psychology of categorization proper, and review how small differences in terms of similarity to a prototype affect decisions of membership to a category.

In the remaining sections we survey various lines of experimental work based on transition series between distinct prototypes. Such series, omnipresent in several domains of experimental psychology, involve so-called morphs, namely gradual alterations of a prototype connecting it to another prototype. Section 3 looks at two paradigms involving unordered presentations of stimuli drawn from such morphing series: the first concerns studies on *categorical perception*, the second concerns studies of the effect of *simultaneous presentation* of stimuli on categorization. Finally, section 4 surveys work on *dynamic sorites*, that is on ordered transitions between prototypes.

To highlight the importance of such transition series in psychology, we deliberately reproduce several examples of stimuli in this chapter. One message of this chapter is that the manner in which such stimuli are presented (whether isolated, in pairs, in random order, or in a specific order) is essential to the way in which similarity between stimuli influences their assignment to a common category.

1. Slippery slopes and the psychology of reasoning

Traditional definitions of a sorites argument distinguish a narrow sense and a broad sense of the term. In the entry "Sorites" of Peirce and Baldwin's *Dictionary of Philosophy and Psychology*, two senses are distinguished in that way (Peirce & Baldwin, 1902). On a specific and marked sense, it is a particular fallacy, namely the sophism of the heap of wheat usually credited to Eubulides of Megara. On a generic and neutral sense, a sorites is merely a "chain of syllogisms".

The two meanings are obviously related, because the sophism of the heap can be presented as such a chain of syllogisms. Le Chevalier de Jaucourt, in the earlier *Encyclopédie* of Diderot and D'Alembert, writes about the argument of the heap: "that argument is composed of several propositions, differing little from one another, and chained in such a way that, after beginning with a manifest and incontrovertible truth, one moves, little by little, to an obviously false conclusion" (cited in Cayrol, 2016). However, not all chains of arguments need be faulty according to the broad definition of a sorites. Le Chevalier de Jaucourt, in the same entry, mentions a number of precautions that one may take in order for a chain of arguments, that is a sorites in the generic sense, to preserve the truth of its first premise down to its final conclusion.[1]

[1]He writes:

> "To avoid surprise, one needs to ensure that everything that is said of the attribute be also said of the subject. That there be no ambiguity in the terms, nor in the propositions. That one insert no negative propositions among affirmative ones. That the proposition that immediately precedes the conclusion not be negative, unless the conclusion might also be negative. That the link and gradation that must be between the propositions be