The questionable utility of "cognitive ability" in explaining cognitive illusions

Ralph Hertwig
Max Planck Institute for Human Development, 14195 Berlin, Germany
hertwig@mpib-berlin.mpg.de
www.mpib-berlin.mpg.de/ABC/Staff/Hertwig

Abstract: The notion of "cognitive ability" leads to paradoxical conclusions when invoked to explain Inhelder and Piaget's research on class inclusion reasoning and research on the inclusion rule in the heuristics-and-biases program. The vague distinction between associative and rule-based reasoning overlooks the human capacity for semantic and pragmatic inferences, and consequently, makes intelligent inferences look like reasoning errors.

Why do most undergraduates appear to get the Linda problem wrong? After all, this problem is meant to instantiate the *inclusion rule,* "perhaps the simplest and most fundamental principle of probability. . . . If A includes B then the probability of B cannot exceed the probability of A" (Kahneman & Tversky 1996, p. 585). Stanovich & West (S&W) (1998b, p. 307) argue that although the problem tests reasoning in accord with a simple rule, "correct responding on the Linda problem . . . is associated with higher cognitive ability." The finding that higher SAT scores are correlated with inclusion responses in the Linda problem is a flagship example of their more general claim that there are two reasoning systems, one associative and the other rule-based, and that students with higher cognitive ability are more likely to give rule-based responses. In what follows, I demonstrate that S&W's use of cognitive ability to explain violations of the inclusion rule, when viewed in light of other findings on reasoning about class inclusion, gives rise to paradoxical conclusions.

Is the cognitive ability of eight-year-olds higher than that of undergraduates? In their classic book *The early growth of logic in the child,* Inhelder and Piaget (1964, p. 101) reported an experiment in which they showed five- to ten-year-old children 20 pictures, four representing colored objects and 16 representing flowers. Eight of the 16 flowers were primulas, four yellow and four of other colors. The children were asked a list of questions about class inclusion relations, one of which was: "Are there more flowers or more primulas?" Only 47% of the five- to seven-year-olds gave answers in accord with class inclusion, that is, which reflected an understanding that the flowers were more numerous than the primulas. Among eight-year-olds, however, a majority (82%) gave responses consistent with class inclusion. Inhelder and Piaget (1964) concluded that "this kind of thinking is not peculiar to professional logicians since the children themselves apply it with confidence when they reach the operational level" (p. 117).

A couple of decades later, Tversky and Kahneman (1983) gave undergraduates at universities such as Stanford the description of a person, Linda, and asked them to rank statements about Linda according to their *probability.* Among them were Linda "is a bank teller" (T) and "is a bank teller and is active in the feminist movement" (T&F). Only 11% of the adult participants ranked T as more probable than T&F, although T&F is included in T. Here we encounter the puzzle. The Linda problem is analogous to the flower problem in that both represent an inclusion relation (Reyna 1991, p. 319). Why, then, do children as young as eight (or nine and ten; Inhelder and Piaget were probably too optimistic about the onset of class-inclusion reasoning; Reyna 1991) follow class inclusion, while undergraduates do not? To the extent that "correct" responding in inclusion problems is associated with higher cognitive ability, as S&W's account suggests, we ought to conclude that eight-year olds have higher cognitive ability than Stanford undergraduates. Not according to Piaget's theory of cognitive development, or for that matter, according to probably any other theory of cognitive development, much less according to common sense: concrete-operational children should trail far behind the undergraduates, who have reached the highest state of cognitive ability, the formal-operatonal stage.

Is the cognitive ability of second graders higher than that of sixth graders? Perhaps the cognitive ability explanation would lead to less paradoxical conclusions if applied only to studies using the Linda and similar problems. In a pertinent study, Davidson (1995) gave second, fourth, and sixth graders problems such as the Mrs. Hill problem. Mrs. Hill "is not in the best health and she has to wear glasses to see. Her hair is gray and she has wrinkles. She walks kind of hunched over." Then, the children were asked to judge how likely Mrs. Hill was to have various occupations, such as Mrs. Hill is "an old person who has grandchildren," and "an old person who has grandchildren and is a waitress at a local restaurant." In Davidson's study, second graders gave more class inclusion responses than sixth graders (65% vs. 43%). Why? If "correct" responding in Linda-type problems is in fact associated with higher cognitive ability, then we ought to conclude that second graders have higher cognitive ability than sixth graders. Again, on any account of cognitive development and common sense, this conclusion is implausible. Ironically, Davidson (1995) interpreted the finding as evidence that children with higher cognitive ability (older children) are more likely to use the representativeness heuristic than children with lower cognitive ability (younger children). Yet the representativeness heuristic seems to epitomize what S&W refer to as the associative system, and thus its use should be correlated with lower cognitive ability.

Why do people violate the principle of class inclusion in the Linda problem? Is there a way out of these paradoxes? In my view, notions such as "cognitive ability," which explain everything and nothing, will not be of much help to us in understanding people's reasoning abilities; nor will "dual-process theories of reasoning," unless underlying cognitive processes are clearly specified (for a critique of such theories, see Gigerenzer & Regier 1996). Problems such as Linda, the cab problem, and the standard Wason selection task are inherently ambiguous (e.g., Birnbaum 1983; Hilton 1995). The Linda problem, for instance, is not a mere instantiation of the inclusion rule. It is laden with the ambiguity of natural language. Take the word "probability." In probabilistic reasoning studies, "probability" is typically assumed to be immediately translatable into mathematical probability. From its conception, however, "probability" has had more than one meaning (e.g., Shapiro 1983), and many of its meanings in contemporary natural language have little, if anything, to do with mathematical probability (see Hertwig & Gigerenzer 1999). Faced with multiple possible meanings, participants must infer what experimenters mean when they use the term in problems such as Linda. Not surprisingly, participants usually infer nonmathematical meanings (e.g., possibility, believability, credibility) because the Linda problem is constructed so that the conversational maxim of relevance renders the mathematical interpretation of "probability" implausible (Hertwig & Gigerenzer 1999). The failure to recognize the human capability for semantic and pragmatic inferences, still unmatched by any computer program, can lead researchers to misclassify such intelligent inferences as reasoning errors.

In contrast to the probability instruction in the Linda problem, Inhelder and Piaget asked children whether there are "more flowers or more primulas." "More" refers directly to numerosity and does not leave open as many possible interpretations as the semantically ambiguous term "probability." Similarly, asking for "frequency" judgments in the Linda problem avoids the ambiguity of "probability" by narrowing down the spectrum of possible interpretations. This is a crucial reason why frequency representations can make conjunction effects disappear (Hertwig & Gigerenzer 1999; for another key reason, namely, the response mode, see Hertwig & Chase 1998).

In sum, Stanovich & West present the Linda problem as support for their thesis that higher cognitive ability underlies correct judgments in reasoning tasks. Whether applied to research by Inhelder and Piaget or to research within the tradition of the heuristics-and-biases program, however, the notion of "cognitive ability" gives rise to paradoxical conclusions. Rather than resort to ill-specified terms and vague dichotomies, we need to analyze cognitive processes – for instance, application of Gricean norms of conversation to the task of interpreting semantically ambiguous terms – which underlie people's understanding of the ambiguous reasoning tasks. Otherwise, *intelligent* inferences will continue to be mistaken for reasoning errors.