4?

Journal of Personality and Social Psychology
1993, Vol. 65, No. 6, 1119–1131

0022-3514/93/$3.00

How Alike Is It Versus How Likely Is It: A Disjunction Fallacy in Probability Judgments

Maya Bar-Hillel and Efrat Neter

One event cannot be more probable than another that includes it. Judging P(A & B) to be higher than P(A) has been called the *conjunction fallacy.* This study examined a *disjunction fallacy.* Ss received brief case descriptions and ordered 7 categories according to 1 of 4 criteria: (a) probability of membership, (b) willingness to bet on membership, (c) inclination to predict membership, and (d) suitability for membership. The list included nested pairs of categories (e.g., Brazil–South America). Ranking a category more probable than its superordinate, or betting on it rather than its superordinate, is fallacious. Prediction, however, may be guided by maximizing informativeness, and suitability need conform to no formal rule. Hence, for these 2 criteria, such a ranking pattern is not fallacious. Yet ranking of categories higher than their superordinates was equally common on all 4 criteria. The results support representativeness against alternative interpretations.

The extension rule in probability theory states that if A is a subset of B, then the probability of A cannot exceed that of B. A special case of the extension rule is the conjunction rule, which states that the probability of A&B can exceed the probability of neither A nor B, since it is contained in both.

Tversky and Kahneman (1983) demonstrated that, under certain circumstances, people predictably and systematically violate the conjunction rule. In one study, they gave subjects the fol ng description:

> ıda is 31 years old, single, outspoken, and very bright. She ma- ed in philosophy. As a student, she was deeply concerned with ıes of discrimination and social justice, and also participated in :i-nuclear demonstrations. (p. 297)

This was followed by a list of eight possible outcomes, each describing possible activities of Linda at the present time (her job, her interests, or both). Subjects were asked to rank order the outcomes by the probability that they describe Linda's current activities. Of the eight, one was representative of Linda ("Linda is active in the feminist movement"), one was unrepresentative of Linda ("Linda is a bank teller"), and one was a conjunction of these two ("Linda is a bank teller and is active in the feminist movement"). A large majority of the subjects (85%) rated the conjunctive outcome, "Linda is a bank teller and is active in the feminist movement," more probable than "Linda is a bank teller."

This result was predicted from the representativeness hypothesis: "Representativeness is an assessment of the degree of correspondence between a sample and a population, an instance and a category, an act and an actor or, more generally, between an outcome and a model" (Tversky & Kahneman, 1983, p. 295). Kahneman and Tversky (1972, 1973) provided much evidence that people often judge the probability of an outcome given a model by the extent to which the outcome represents the model. In addition, Tversky (1977) showed that adding to an outcome (O) a feature (F) that matches a model (M) enhances the match between the outcome and the model. In other words, the match of O&F to M can be greater than the match of O to M. Hence, insofar as people judge the probability of outcomes by their representativeness, being a bank teller and active in the feminist movement would be judged more likely an outcome for Linda than being a bank teller, due to the addition of a feature that is representative of Linda (feminism) to her unrepresentative job. Whereas there is nothing logically wrong with the judgment that being a feminist bank teller is more *representative* of Linda than being a bank teller, judging the conjunctive outcome to be more *probable* than its constituent violates the logically necessary conjunction rule.

Another special case of the extension rule is the disjunction rule, according to which the probability of A-or-B can be smaller than neither the probability of A nor the probability of B, since it contains both. Formally speaking, there is no difference between the three rules (conjunction, disjunction, and extension), because for any pair of events A and B in which B is a subset of A, A can always be represented as a disjunction, one of whose constituents is B, and B can always be represented as a conjunction, one of whose constituents is A. For example, one can argue that the set of bank tellers is a disjunction—of bank tellers who are active feminists with bank tellers who are not. Viewed in this way, Tversky and Kahneman's (1983) results could just as well have been labeled the *disjunction fallacy.* Why then are they regarded as a *conjunction fallacy?* Is this just a matter of arbitrary choice?

Formally speaking, the answer is yes, but psychological considerations favor one view over another. Consider the category *parent* and its subcategory *mother.* One can just as well choose to view *mother* as the conjunction of *parent* with *female* as to

Maya Bar-Hillel and Efrat Neter, Department of Psychology, Hebrew University, Jerusalem, Israel.

Correspondence concerning this article should be addressed to Maya Bar-Hillel, Department of Psychology, Hebrew University, Jerusalem 91905 Israel.

view *parent* as the disjunction of *mother* with *father.* In contrast, the category *bank teller* does not naturally evoke a representation as a union, and certainly not as a union of bank tellers who are active feminists with bank tellers who are not. At the same time, the subcategory of bank tellers who are active feminists can hardly be described except by resort to the conjunction of these two constituents. Indeed, the language does not even contain a single-word label to designate this category. In that sense, the categories *bank teller* and *bank teller and active feminist* are more naturally viewed, respectively, as a unitary category and a conjunction of two categories, than as a disjunction of two categories and a unitary category.

How, then, might one create a category that would be naturally viewed as disjunctive? The simplest possibility to come to mind is to replace the connective *and* used to create conjunctive categories with the connective *or.* This idea must be implemented with caution, however, because the English words *and* and *or* do not always quite correspond to the logical connectives *and* and *or.* First, the English *or* is often understood in its exclusive sense of "A or B but not both," as in "The party will take place next week or the week after." Second, the English *and* can be used to create a union as well as an intersection—the sentences "She invited colleagues *or* relatives" and "She invited colleagues *and* relatives" could be used to describe the same guest list. Third, and most pertinent to present concerns, not all categories that can be meaningfully joined by one of these connectives lend themselves to as meaningful a joining by the other. For example, whereas putting *and* between *bank teller* and *active in the feminist movement* creates a meaningful category, putting *or* between these two category names creates a rather odd one. Similarly, whereas the question, "Is Linda more likely to be a bank teller, or a bank teller and active in the feminist movement?" makes some sense, the question, "Is Linda more likely to be a bank teller, or a bank teller or active in the feminist movement?" sounds to us rather confusing.

Nonetheless, this was precisely the approach taken by all previous attempts to study a disjunction fallacy. In the first attempt to extend the conjunction fallacy to a disjunction fallacy, Morier and Borgida (1984) gave subjects Linda's description and asked them to estimate the probability that (a) Linda is a bank teller, (b) Linda is active in the feminist movement, (c) Linda is a bank teller and is active in the feminist movement, and (d) Linda is a bank teller or is active in the feminist movement.

Wells (1985) took a similar approach. He also gave subjects personality descriptions and asked them for the probabilities of four events: two individual events, their conjunction, and their disjunction. Wells, however, took two precautions to make the resulting disjunction more natural. First, both of the individual events were attitudes toward some national issues, making them of a kind that sounds quite natural when joined by *or.* Second, Wells explicitly added *or both,* to highlight that the *or* was nonexclusive. For example, after describing Jim, subjects were asked for the probability that Jim (a) "favors . . . a U.S. buildup in military strength," (b) "favors the decriminalization of marijuana," (c) "favors [the first] *and* [the second]," and (d) "favors [the first], or [the second], or both." Yet, in spite of these precautions, Wells himself remained concerned "that these subjects misunderstood the union request (e.g., interpreted "or" as a conjunction)" (p. 277).

Biela (1986) took a slightly less formal approach when creating conjunctive and disjunctive categories. After describing to physicians one or two symptoms of an otherwise unknown patient, he asked them to attach degrees of confidence to each of the following diagnoses: (a) "I would predict hypertensive encephalopathy," (b) "I would predict hypertensive retinopathy," (c) "I would predict that either hypertensive encephalopathy or hypertensive retinopathy (or both) is the case," and (d) "I would predict both hypertensive encephalopathy and retinopathy." The subtle variation in formulation between the intended disjunctive (c) and conjunctive (d) diagnosis show Biela's awareness that mechanical joining of the two diagnoses by *or* or *and* is unsatisfactory. Still, (d) might well be understood by a reader to mean "I would give the following degree of confidence to *either* of these diagnoses." In the context of (a), (b), and (c) this possibility becomes less likely, but this comment is intended to show how tricky it is to create categories that are unambiguously disjunctive.

Another difference between categories formed by the connective *or* and categories formed by *and* lies in the difference between matching an instance to A-and-B versus to A-or-B. To be concrete, consider matching Linda to the conjunction "feminist and bank teller," on the one hand, and to the disjunction "feminist or bank teller," on the other. The conjunction requires a single comparison, albeit to a compound event. The disjunction, however, seems to require two comparisons, one for each of the constituent events. Tversky's (1977) theory relating similarity judgments to stimuli's features makes no predictions concerning how two constituent similarity judgments are combined to yield a single similarity judgment for the disjunctive event.

Judging the similarity of the compound event can be sidestepped, however, as it was in Carlson and Yates' (1989) study. Subjects were presented with pairs of events, such as "Syria and Israel will sign a peace treaty by the end of this year" and "The Bill Cosby show will not be one of the top 10 rated TV shows at the end of the season," as well as the conjunction and the disjunction of these events, and they were asked to rank them by their probability of actually occuring within the coming year. But because no model was provided according to which subjects could judge representativeness, the probability of the compound events could only be derived from some kind of combination of the probabilities of its constituent events, not from similarity judgments.

The type of problems used by Carlson and Yates (1989) were termed "probability combination problems" by Gavanski and Roskos-Ewoldsen (1991). These authors recently showed that, when constituent probabilities were controlled for, the rates of the conjunction fallacy were similar in problems in which representativeness could be used to assess the probability of the conjunctive event (e.g., the Linda problem) and in problems in which representativeness could not possibly be used (e.g., probability combination problems). This led them to conclude that the conjunction fallacies "stem primarily from the incorrect rules people use to combine probabilities" and that the "only contribution of representativeness stems from its effects on a conjunction's component events" (p. 190). Insofar as the disjunctive events used in all studies to date were created by joining constituent events by the connective *or*, the same conclusion could apply to the disjunction fallacy.

The purpose of the present study was twofold: First, to explore the possibility of extending the conjunction fallacy to a more general extension fallacy, while using natural disjunctive categories rather than the somewhat mechanically derived ones used by previous researchers; Second, to explore whether extension fallacies can be obtained even in problems that involve no compound events, hence the fallacy could not possibly result from incorrect combination rules.

Study 1

Method

For this study, we sought categories that would be disjunctive in character, yet defined without the connective *or*.[1] Our solution derived from Rosch's work on categorization (e.g., Rosch, 1978). Recall that the classification of objects into categories is not unique. Thus, the same object can be classified as "artifact," "furniture," "chair," "office chair," or "Eleanor's black leather office chair," to mention but a few possibilities. Yet most people would usually think of it as just "chair." In the hierarchy of nested categories, the level at which objects are commonly classified is known as "the basic level category" (Rosch, 1978). The basic level is not determined arbitrarily. Objects at this level have many features in common—considerably more than those shared by members of higher level categories, but only slightly less than those shared by members of lower level categories. For example, the number of attributes commonly seen as common to *furniture* is 3, to *chair* is 9, and to *kitchen chair* is 10 (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976, Table 2, p. 388).

We capitalized on the observation that basic level categories are typically unitary (i.e., standardly thought of as neither conjunctive nor disjunctive categories), whereas higher levels are often disjunctive (e.g., "furniture" is the union set of "chairs," "tables," "beds," "cupboards," etc.), and lower levels are often conjunctive (e.g., "leather chair"). However, superordinate categories often enjoy a name or label of their own, and are not explicitly defined as compound sets. If a described instance were to match a unitary category more than it matched its superordinate, a disjunction error could potentially arise. Moreover, in the absence of explicit compounding, such error could not be attributed to incorrect combination rules.

Table 1 shows the five types of questions that we devised: (a) The instance is a brief personality description of a student, and the options are fields of university study (Danielle and Oded); (b) The instance is a brief personality description, and the options are places of residence (Gidi and Eldar); (c) The instance is a description of a place, and the options are countries or continents (Gila and Na'ama); (d) The instance is a brief personality description, and the options are social-political movements (Eli and Ze'ev); and (e) The instance is a brief personality description, and the options are causes of death (Yaron and Y.C.).

The target option (i.e., the one designed to be most representative of the instance) was selected to be a basic level category name. The superordinate categories are natural ones, designated by familiar proper names or labels. To strengthen the tendency to view them as a union, their names were prefaced by the words *one of the* or *some kind of*. For example, the basic level cause of death "car accident" was paired with the superordinate "some kind of unnatural cause."

All questions were accompanied by seven options. To give a typical example, the list of fields of university studies provided with the description of Danielle was literature, humanities, physics, natural sciences, geography or geology, statistics, and political science (not in this order). It is important to note that the lists always included two pairs of a basic level category and its superordinate category (here, the first four options).[2] Also, there was an option defined with the connective *or* (here, the fifth option). The actual ordering of the various options varied from problem to problem.

Israeli university students taking summer courses (about 60 in advanced economics, 40 in elementary economics, 40 in statistics, 60 in mathematics, 35 in accounting, 15 in sociology) were given from one to four prediction problems, but never two of the same type. On the whole, this subject population is comparable to those identified in other studies (e.g., Tversky & Kahneman, 1983; Wolford, Taylor, & Beck, 1990) as "informed," although subject sophistication was not one of the variables studied here.

There are several concerns a disjunction fallacy might raise. Although some of them were previously raised—and answered—by Tversky and Kahneman (1983) with regard to the conjunction fallacy, in the present context the concerns take on somewhat different twists, and we used somewhat different approaches to deal with them.

First, we were concerned that subjects who see a disjunctive category listed alongside one of its constituents would interpret the disjunctive category in the subtractive sense. For example, given "one of the European countries" alongside "Switzerland," some subjects might implicitly exclude Switzerland from the European countries. Even though Tversky and Kahneman (1983) and Morier and Borgida (1984) ruled out this account of the conjunction fallacy, and it seems unlikely altogether with our natural categories, we thought it prudent to explicitly discourage the possibility.

Tversky and Kahneman (1983) handled this concern by way of what they called "the indirect test" (p. 297). Some subjects were given a list of outcomes that included either the larger category or its subcategory, but not both. When the larger category is not listed alongside one of its subcategories, there is no reason to interpret it as excluding that subcategory. In an indirect test it is impossible to rank "bank teller" before "bank teller and active feminist," because no list includes both options. But across groups, it is possible for "bank teller" to receive a mean rank that is higher than that given to "bank teller and active feminist."

Our approach was simply to include in the subjects' instructions the following explicit cautionary note:

> It is quite possible for something to be included in more than one option listed. For example, suppose the list is of foodstuffs, and one option is "frozen foods", while another is "desserts". In this case, do not interpret "frozen foods" as "frozen foods excluding desserts", nor interpret "desserts" as "desserts which are not frozen"—"ice-cream" qualifies as both "frozen food" and "dessert".

The second concern has to do with linguistic conventions. In their discussion of the conjunction fallacy, Tversky and Kahneman brought up the question of how the Gricean concept of cooperativeness (Grice, 1975) applies to a speaker who is less than certain of the truth of his or her assertions. They suggested that such a speaker "may wish to follow the maxim of value: Select the message that has the highest expected value" (Tversky & Kahneman, 1983, p. 312), where "the expected value of a message can be defined as its information value if it is true, weighted by the probability that it is true" (p. 312). With respect to the assertion

---

[1] Study 1 began as a class project. We gratefully acknowledge the contribution of Anat Ben-Simon, Eyal Gamliel, and Hannan Goldschmidt to the study. We thank Danny Kahneman and Ilan Yaniv for helpful suggestions.

[2] Actually, there are two exceptions—Hadar ha'Carmel, which is a Haifa neighborhood (Gidi), and North Tel Aviv, which is a Tel Aviv neighborhood (Eldar), but these were not target pairs.

Table 1
Stimuli Used in Experiment 1

Character	Target category	Target superordinate category	Foil category	Foil superordinate category	Description
Danielle	Literature	Humanities	Physics	Natural sciences	Sensitive and introspective. In high school she wrote poetry secretly. Did her military service as a teacher. Though beautiful, she has little social life, since she prefers to spend her time reading quietly at home rather than partying. What does she study?
Oded	Physics	Natural sciences	Literature	Humanities	Did his military service as a combat pilot. Was a brilliant high school student, whose teachers predicted for him an academic career. Independent and original, diligent and honest. His hobbies are shortwave radio and Astronomy. What does he study?
Gidi	Tel Aviv	Dan Metropolitan Area	Hadar ha'Carmel	The North of Israel	23 years old, he wears the latest fashions, and drives a new sportscar. He spends time in discotheques and expensive pubs, and is a social butterfly. He is occasionally mentioned in the gossip columns. Where does he live?
Eldar	Kibbutz Rosh Ha'Nikra	The North of Israel	North Tel Aviv	Dan Metropolitan Area	23 years old, he dresses modestly, and wears sandals year-round, even in winter. He is tanned from outdoor work. He spends his leisure time hiking in the countryside. Where does he live?
Gila	Japan	Asia	Canada	North America	Writes letter home describing a densely populated country, in which modern technology coexists with an ancient local culture. The people, hardworking and competitive, are not inclined to invite strangers into their homes. Where was the letter written?
Na'ama	Switzerland	Europe	Brazil	Latin America	Writes letter home describing a country with snowy wild mountains, clean streets, and flower decked porches. Where was the letter written?
Eli	Peace Now	A peace movement	Gush Emunim	A national movement	39 years old, a professor of Greek Philosophy and Ethics, he holds socialist views. Following the Lebanon War he became politically active, while remaining a "bleeding heart." Where is he active?
Ze'ev	Gush Emunim	A national movement	Peace Now	A peace movement	39 years old, with a red beard and fiery eyes. Married, and a devoted father of five. He teaches Holy Studies in a West Bank settlement. He immigrated from the US 18 years ago. He spends his leisure time hiking in the countryside. Where is he active?
Yaron	Road accident	Unnatural cause	Cancer	Disease	Till he died suddenly at 27, was fully of gaity and life. A womanizer, he often drank, and acted wildly in parties and on the road. To his concerned friends he always said, "It couldn't happen to me." What did he die of?
Y.C.	Lung cancer	Cancer	Road accident	Unnatural cause	A woman who smoked over a packet a day for over 10 years. What did she die of?

of disjunctions, it is "*misleading* (though not false) for a speaker to make a disjunctive assertion where he is in a position to assert one of the disjuncts. Thus, when a speaker asserts A or B he cannot assert A by itself since it is too uncertain" (Adams, 1975, pp. 19–20). For example, I am uncertain when I saw you last, and I assert that "it was either at last year's Psychonomics or when I was in Boston this summer." A listener who assumes Gricean cooperativeness would be justified in inferring from this assertion that I believe the disjunction is fairly likely, but neither of its constituents is sufficiently likely to be asserted by itself.

Applying the value maxim to our problems, a cooperative speaker under uncertainty would only assert, "Danielle is majoring in the humanities" if he or she is not in a position to assert the more informative "Danielle is a literature major." If, on the other hand, one can assert, "Danielle is a literature major," one would not assert, "Danielle is majoring in the humanities," because that would mislead a listener into thinking, counterfactually, that one could not have asserted, "Danielle is a literature major." Hence, the disjunctive assertion, rather than following logically from the more specific assertion, is actually incompatible with it, in the sense that the two would be made under nonoverlapping conditions. The rules of conversational implicature differ from those of logical entailment.

In a betting paradigm (as well as in some legal contexts, most notably contracts), however, logical entailment overrides conversational implicatures. The proposition that "Danielle is majoring in the humanities" is undeniably verified upon learning that "Danielle is a literature major." Even a cooperative speaker, who in everyday conversation would not assert that "Danielle is majoring in the humanities" when "Danielle is a literature major" could be asserted, can legitimately prefer to bet that "Danielle is majoring in the humanities" than that "Danielle is a literature major." The conditions for asserting propositions do not coincide with those for betting on them. In particular, one might sometimes legitimately prefer to assert A than B, yet prefer to bet on B than on A.

To address this concern, we ran two variants of our problems. In the first, subjects were asked to rank the listed categories by the probability that they included the described case as a member. In the second, a different group of subjects ranked the categories by their willingness to bet that the described case was a member in these categories. In this task, no mention whatsoever was made of the word *probability* (thus excluding any confusions this term might engender). In a betting paradigm, the question of why particular outcomes were chosen to be stated and not others is not as relevant as it is for the original task. There may be conditions under which one would be more likely to assert A than A-or-B, but there are no conditions under which a bet on A is more likely to win than a bet on A-or-B.

Results and Discussion

The categories used in Study 1 were not studied in the categorization literature (which typically uses object categories). Hence, a manipulation check was in order, to establish that our target categories really are basic level ones, in the sense that they are at the level where people would tend to locate instances. We asked a separate group of 48 subjects to answer the following questions: (a) What does your best friend study? (100%); (b) Where do your parents live?[3] (80%); and (c) To what place in the world would you most like to travel? (75%). As hoped, most respondents answered the questions at the level of the target category, namely they gave answers like (a) physics (rather than natural sciences or quantum theory), (b) Haifa (rather than up north or Hadar ha'Carmel), (c) Japan (rather than the Far East or Tokyo), respectively. Note the shift in location of the basic level that occurs between the questions "Where do your parents live?" and "Where would you like to travel?," for obvious contextual reasons.

The percentages in parentheses after the questions above indicate how many respondents answered each question at the expected level. In the first question, 11 said, "Does not study at all," so 100% is based on the 37 who responded; in the second question, three people said in general "city" or "in the country," and the rest added a neighborhood name to the town name; in the third question, one person gave a state ("Hawaii"), three people gave city names (e.g., "Moscow"), and the rest gave either just continents (e.g., "Africa") or continents alongside country names (e.g., "South America [Argentina]").

We also checked that our subjects knew that our target categories were contained in their superordinate categories. The same 48 subjects also answered the following questions: (a) In which faculty is the Department of Physics? (96%); (b) In which faculty is the Department of Literature? (98%); and (c) What part of the country is Rosh Ha'Nikra in? (98%). No subjects answered these questions in error, although one or two did not answer at all.

The other categories were not checked (e.g., we did not check that "car accident" is a basic level category whereas "unnatural cause" is its superordinate category, nor that people know that car accidents are unnatural deaths) because the check seemed either superfluous or tricky (see, however, the manipulation check in the following Study 2).

The results of the disjunction study are shown in Table 2. The top row shows the rates of violation of the disjunction rule for the "rank the outcomes by your willingness-to-bet on them" formulation, the middle row shows the rates for the "rank the outcomes by their probability" formulation, and the bottom row combines the two conditions, which of course are normatively equivalent. (The rates in parentheses will be explained later.)

There are small but systematic differences between the probability version and willingness-to-bet version. The rate of the disjunction fallacy is lower under the betting formulation in 8 of the 10 cases (the exceptions being Gidi and Y.C.). However, the decrease averages less than 10% (even the most extreme difference, obtained for Eldar, and the overall difference, fall just short of significance, $z = 1.57$, $z = 1.65$, respectively). The rates of the disjunction fallacy never fall below 50% in either task. We conclude that the disjunction fallacy cannot be accounted for only by conversational implicatures.

The rate of the disjunction fallacy in our problems averaged .64, and never exceeded .83 (Gidi). This is less than the typical rates reported for the conjunction fallacy. Why might this be the case? Perhaps the disjunction rule simply is somehow more compelling than the conjunction rule, although this explanation has a distinct ad hoc flavor.[4] More likely, note that Tversky

[3] We did not ask questions (a) and (b) about the respondents themselves, because all those sitting in a class together at the time of answering the questionnaire studied the same topic and lived in the same city.

[4] We hasten to note that in both Study 1 and Study 2 we did not select stimuli for presentation on the basis of their effectiveness in eliciting the disjunction fallacy, so the readers are not seeing a selective set of best shots. We did throw out some results, however, as in one problem in

Table 2
Disjunction Fallacy Rates for the Target (Foil) Pairs in Experiment 1

Condition	Danielle	Oded	Gidi	Eldar	Gila	Na'ama	Eli	Ze'ev	Yaron	Y.C.	Overall
Betting	50 (18)	65 (24)	74 (77)	61 (3)	68 (23)	71 (32)	57 (30)	63 (29)	71 (39)	52 (54)	61 (32)
N	66	66	35	31	40	31	63	51	28	65	476
Probability	56 (31)	67 (27)	83 (51)	79 (15)	78 (40)	79 (24)	57 (28)	68 (41)	69 (37)	58 (61)	66 (37)
N	68	70	35	33	36	33	85	56	29	90	535
Combined	53 (25)	66 (26)	79 (64)	70 (9)	73 (31)	75 (28)	57 (29)	66 (35)	70 (38)	55 (58)	64 (35)
N	134	132	70	64	76	64	148	107	58	155	1,011

and Kahneman's (1983) inclusive categories were highly nonrepresentative (e.g., "bank teller" as an occupation for Linda), whereas the representativeness of the conjunctive category was much enhanced by conjoining a representative feature (e.g., being an active feminist). In contrast, both levels of our target categories were representative. For example, Oded sounds much like a physics student, but because physics is a typical department in the natural sciences, he also sounds much like a student in the natural sciences (this is demonstrated explicitly later in Study 2).

In general, Table 3 shows the mean rank received by the superordinate category and its subset, as well as the rank of these means. The target category always received the lowest mean rank, and its superordinate category received the next lowest mean rank.[5] So it appears that the target pair of nested categories consisted of the two most representative categories in the list, which entails an obvious loss of power.

If a basic level category and its superordinate are both representative of some description, why does the subcategory rank as more probable? Perhaps this is a mere reflection of people's predisposition to answer categorization questions at a basic level. To test this possibility, each list in our study contained, in addition to the target category and its superordinate disjunctive category, a foil basic level category accompanied by its superordinate disjunctive category. The foil pair was designed to be unrepresentative of the instance (e.g., physics and natural sciences for Danielle, literature and humanities for Oded; see Table 1). We can compare the disjunction rate for these pairs to the disjunction rate for the target pair.

The parenthetical numbers in Table 2 show the disjunction fallacy rates for the foil pairs alongside those for the target pairs. In 8 out of the 10 descriptions (the exceptions are Gidi and Y.C.), these rates are smaller for the foil pair than for the target pair, and even smaller than 50%. On average, the rates for the foil pair are just over one half of those for the target pair (35% vs. 64%, respectively). Moreover, in two category types (a and d) we can compare exactly the same two pairs in reversed roles. The pair literature–humanities is target for Danielle and foil for Oded, whereas the pair physics–natural sciences is target for Oded and foil for Danielle; the target–foil pairs are similarly reversed with regard to the Peace Now–peace movement versus Gush Emunim–nationalistic movement for Eli and Ze'ev. The disjunction fallacy rate is about twice as high in each of these pairs when they serve as target pairs as when the same pair is a foil pair. (For the probability results: .56 vs. .27, $z = 3.46$; .67 vs. .31, $z = 4.23$; .57 vs. .41, $z = 1.77$; .68 vs. .28, $z = 4.68$. For the betting results: .50 vs. .24, $z = 3.09$; .65 vs. .18, $z = 5.48$; .57 vs. .29, $z = 2.99$; .63 vs. .30, $z = 3.52$.)

The preference for predicting a basic level category over its superordinate category clearly interacts with whether the pair of nested categories is representative or not. When a category is a highly likely one, people seem to prefer the narrower possibility ("Oded studies physics") over the broader one ("Oded studies one of the natural sciences"). However, when the category is an unlikely one, they prefer the broader possibility ("Danielle studies one of the natural sciences") to the narrower one ("Danielle studies physics").

This is compatible with hypotheses put forth by Tversky (1977). Recall that, according to Rosch (e.g., 1978), superordinate categories (e.g., furniture) have fewer features in common than basic level ones (e.g., chairs). Recall also that the addition of similar features increases similarity, whereas the addition of dissimilar features decreases it. Thus, physics, being richer in details that are compatible with Oded's description, is more like Oded than is natural sciences, whereas literature, being richer in details that are incompatible with Oded's description, is less like Oded than is humanities. The probability ranking, quite in accordance with the representativeness hypothesis, follows the similarity ranking. In Study 2, we subjected this hypothesis to a direct test.

Study 2

Method

The second study had three purposes. In ascending order of importance these were (a) to extend and replicate the earlier study with new and more powerful stimuli; (b) to explore, in addition to the criteria of probability and willingness-to-bet, a third criterion for ranking options that is related to these two but is not normatively obligated to coincide with them; and (c) to directly test some of the speculations set forth to account for the results of the first study. We consider these in turn.

which the manipulation check showed that 39% of our subjects were not aware that the basic level target category (geology) was a subset of its superordinate category (natural sciences). This problem actually elicited a high rate of the disjunction fallacy, but it could hardly be called that if people were unaware of the set inclusion.

[5] Although Tversky and Kahneman (1983) did not report the rank order of the mean ranking of their inclusive and conjunctive categories, a look at their Table 1 (p. 298) suggests that such was not the case in their study.

Table 3
Means of Ranks of the Category Pairs Used in Experiment 1 and Their Ranks by Story and Task (Target Categories Are Outside Parentheses, Foil Categories Are Inside Parentheses)

		Means of ranks		Ranks of mean ranks	
Story	Task	Unitary (Basic level)	Disjunctive (Superordinate)	Unitary (Basic level)	Disjunctive (Superordinate)
Danielle	Betting	2.0 (5.8)	2.1 (4.2)	1 (7)	2 (4)
	Probability	2.3 (5.2)	2.5 (4.3)	1 (6)	2 (3)
Oded	Betting	1.5 (6.5)	2.4 (5.6)	1 (7)	2 (6)
	Probability	1.8 (6.3)	2.4 (5.4)	1 (7)	2 (6)
Gidi	Betting	1.4 (4.2)	2.3 (5.8)	1 (4)	2 (6)
	Probability	1.2 (3.9)	2.4 (5.8)	1 (3)	2 (6)
Eldar	Betting	2.1 (6.4)	2.6 (4.3)	1 (7)	2 (5)
	Probability	1.9 (6.0)	3.1 (4.7)	1 (7)	2 (6)
Gila	Betting	1.9 (5.5)	2.4 (4.8)	1 (7)	2 (5)
	Probability	2.1 (5.0)	3.0 (4.6)	1 (6)	2 (5)
Na'ama	Betting	1.4 (3.2)	2.3 (5.6)	1 (3)	2 (6)
	Probability	1.9 (3.4)	2.7 (4.8)	1 (3)	2 (5)
Eli	Betting	2.5 (4.6)	2.6 (4.9)	1 (4)	2 (6)
	Probability	2.2 (5.7)	2.3 (4.9)	1 (7)	2 (6)
Ze'ev	Betting	2.0 (6.3)	2.3 (5.7)	1 (7)	2 (6)
	Probability	2.3 (5.6)	3.1 (5.3)	1 (7)	2 (6)
Yaron	Betting	1.8 (4.8)	3.4 (4.2)	1 (6)	2 (4)
	Probability	1.5 (5.1)	2.9 (4.6)	1 (7)	2 (4)
Y.C.	Betting	2.6 (4.1)	2.6 (4.4)	1 (4)	1 (5)
	Probability	2.5 (3.9)	2.9 (4.8)	1 (4)	2 (5)

Increasing power. In Study 2 we used five new stimuli, shown in Table 4. (a) The instance is a brief personality description of a student and the options are areas of university study (Yossi). (b) The instance is a brief personality description and the options are sports (Doron). (c) The instance is a description of a place and the options are countries or continents (Dorit and Alon). (d) The instance is a brief personality description and the options are Israeli political parties (Asher). As before, the target option was designed to be a basic level category name and questions were accompanied by seven options, including two pairs of nested categories. A new set of Israeli university students (about 100 in advanced economics, 50 in elementary economics, 120 in social work, and 70 in law school) were given either three or four of these problems. The instructions were as in Study 1.

The main difference between the new set of descriptions and the previous one is that we attempted to choose target category pairs such that, although the basic level one would seem to be a very representative possibility, its superordinate would not seem so. This we did either by picking an unrepresentative subset of its superordinate category (e.g., tennis is an atypical ball game; South Africa is an atypical African country) or by constructing a description that picks on features of the subset that the subset does not share with its superordinate (e.g., mathematical skill is required in statistics much more than in the typical social science; during its brief existence, Meimad was the only dovish and liberal–humanistic religious party in Israel). This attempt was intended to boost the rates of the disjunction fallacy.

A new criterion for ranking. Recall our earlier mention of Tversky and Kahneman's (1983) suggestion that cooperative speakers "may wish to follow the maxim of value: Select the message that has the highest expected value" where "the expected value of a message can be defined as its information value if it is true, weighted by the probability that it is true" (p. 312). We argued there that in a rank by probability or rank by willingness-to-bet task, cooperativeness is beside the point, normatively speaking, and the information value of a statement should be totally subjugated to its probability of being true. However, if the task were to rank the options according to how inclined you would be to predict them, the value maxim is quite relevant. In other words, there is nothing normatively inappropriate in being more inclined to predict that "Danielle is a literature major" than the less informative, even if more likely to be correct, "Danielle is majoring in the humanities."

Just as the betting task was expected to make it harder for people to violate the disjunction rule by suggesting truth conditions and by rewarding correctness alone rather than informativeness or cooperativeness as well, so a prediction task was expected to make it easier for them to violate the rule. Indeed, Yaniv and Foster (1990) showed that in a numerical prediction task in which subjects could choose at will the "graininess" of numerical estimates (i.e., the size of interval within which they choose to place their estimate), subjects naturally produce predictions that have a probability close to 50% of being correct—not one close to 90% or 95%. Similarly, subjects often evaluate as better others' finely grained estimates over more coarsely grained ones, even if the latter but not the former includes the true answer. For example, given the question, "What amount of money was spent on education by the U.S. federal government in 1987?" and the correct answer, $22.5 billion, 80% of Yaniv and Foster's subjects said that "$18 to $20 billion" was a better estimate than "$20 to $40 billion," although only the latter included the correct answer. Thus, in this study, we added a rank by tendency to predict task to the two tasks of the first study. Each subject ranked according to a single one of these four criteria.

Testing the representativeness hypothesis directly. In Study 1, we observed a simultaneous tendency to rank the basic level category higher than its superordinate when both were likely options (the target pair) and to rank the basic level category lower than its superordinate when both were not (the foil pair). We speculated that because basic level categories are typically richer in common features than their superordinate categories, moving from the latter to the former typically increases similarity if they are similar to a description (by adding common features), and typically decreases similarity if they are dissimilar to the description (by adding features that are distinctive from the descrip-

Table 4
Stimuli Used in Experiment 2

Character	Target category	Target superordinate category	Foil category	Foil superordinate category	Description
Yossi	Statistics	Social sciences	Hebrew language	Humanities	Recently discharged from service in Intelligence. Outstanding high school student, who once won a national math competition. Pedantic and careful, with a good head for numbers. A computer nut. Shy and reserved, a loner. What does he study?
Doron	Tennis	A ball game	Fast walking	A track and field sport	A successful Jerusalem attorney. Colleagues say his whims prevent him from being a team worker, attributing his success to competitiveness and drive. Slim and not tall, he watches his body and is vain. Spends several hours a week on his favorite sport. What sport is that?
Dorit	South Africa	Africa	Holland	Europe	Writes letter home describing an English-speaking country, where a modern Western technological society coexists with primal landscape and free-roaming wildlife. Where was the letter written?
Alon	Yugoslavia	Europe	Canada	North America	Writes letter home describing a multination country in constant strife, not very advanced economically, but outstanding in its beaches. Where was the letter written?
Asher	Meimad	A religious party	Mapam	A left-wing party	Very religious, but yet a liberal. Attended a religious high school, where he started a group for enhancing tolerance and understanding between the religious and the secular. Today he is more into foreign affairs and security issues, defining himself as a dove. What party does he vote for?

tion). This speculation can be tested directly by asking subjects to rank options according to how suitable they seem to be for the described person. This, then, is the fourth criterion by which options were ranked in Study 2. It allowed us to test directly whether the probability rankings, in accordance with the representativeness hypothesis, indeed follow the suitability rankings.

Results and Discussion

As in Study 1, we begin with a manipulation check. In Study 1, we checked whether subjects knew the relevant set inclusions by direct questioning. This turned out to be awkward for some categories (e.g., "what kind of death is a death by car accident?"). So in Study 2, we used the following indirect method. We presented 25 subjects with a series of questions formulated thusly: "You have bet that someone is an X, and it turns out that she is a Y. Have you lost the bet or won it?" The 50-odd pairs of Xs and Ys included our 10 nested categories (a target pair and a foil pair for each of five descriptions). In addition, we confirmed all the set inclusions that were necessary for Study 1, such as that "road accident" is considered "an unnatural death." In this article, we report only questions in which at least 80% of the respondents answered appropriately (the median was more than 90%).

Following Shafir, Smith, and Osherson (1990), we called the ranking of the superordinate category higher than its subset a *disjunction fallacy* when it is normatively inappropriate to rank them so (as in the probability and the willingness-to-bet tasks), and a *disjunction effect* when there is no such normative constraint (as in the suitability and inclination-to-predict tasks). The results of Study 2 are shown in Table 5. The rows show the rates of the disjunction fallacy by the four tasks, as well as by some task combinations—betting plus probability for a disjunction fallacy rate, prediction plus suitability for a disjunction effect rate.

First, the overall rate of the disjunction fallacy is significantly higher here than in Study 1 (.77 vs. .64, $z = 4.49$). We attribute this rise to the fact that the basic category and its superordinate were, on average, farther apart in this study than they had been in Study 1. This can be seen in Table 6, which shows the mean ranks received by the superordinate category and its subset, as well as the rank of these means. Whereas in Study 1 the target superordinate category always received the next to lowest mean rank (see Table 3), here the target superordinate category received the next to lowest mean rank only four times out of 20 (three of them in the case of Asher).

Second, the magnitude of the disjunction fallacy for the foil categories is again consistently lower than .50 (with two exceptions: Yossi under probability and Alon under prediction), and about one half the magnitude of the rate for the target categories (the overall rate is .36 for the foils vs. .74 for the targets).

However, the small but systematic differences between the probability version and willingness-to-bet version, found in Study 1, is no longer evident. Indeed, it seems to have slightly reversed itself—now the betting rates are higher than the probability rates for four of the five stories (the exception is Alon) as well as on average (.79 vs. .75, $z = .90$, *ns*).

Is the disjunction effect rate in the prediction task, in which we argued that it was not fallacious, higher than in either the betting or probability tasks? No systematic pattern emerges: The prediction rates are sometimes higher (Doron, Dorit, Alon) and sometimes lower (Yossi, Asher) than the rates in the betting and probability tasks. The overall disjunction effect rate in the

Table 5
Disjunction Fallacy Rates for Target (Foil) Pairs in Experiment 2, by Task, and by Combined Tasks

Task	Yossi	Doron	Dorit	Alon	Asher	Overall
Betting	87 (47)	81 (19)	80 (37)	74 (43)	70 (21)	79 (34)
N	38	36	35	35	33	177
Probability	84 (61)	78 (16)	74 (43)	79 (32)	59 (41)	75 (39)
N	38	37	35	34	34	178
Prediction	84 (44)	82 (17)	83 (34)	82 (55)	55 (45)	77 (39)
N	37	34	35	33	33	172
Suitability	85 (49)	67 (6)	83 (33)	73 (35)	50 (33)	68 (32)
N	39	36	40	41	76	232
Disjunction fallacy	86 (54)	79 (18)	77 (40)	77 (38)	64 (31)	77 (36)
N	76	73	70	69	67	355
Disjunction effect	84 (47)	74 (11)	83 (34)	77 (44)	51 (37)	72 (35)
N	76	70	75	74	109	404

prediction task is .77, exactly like the overall disjunction fallacy rate (of the combined probability and betting tasks).

What about the suitability results? In the suitability task there is no normative reason whatsoever to rank superordinate categories higher (or, for that matter, lower) than their subsets. Thus, we might expect the disjunction effect to be largest in this task. Actually, however, it was most often (Doron, Alon, Asher) smallest, and on average (.68) smallest.

Finally, we combined the two conditions in which the disjunction rule is normatively necessary (betting and probability) and the two conditions in which it is not (prediction and suitability) and compared the combined rates. Contrary to norma-

Table 6
Means of Ranks of the Category Pairs Used in Experiment 2 and Their Ranks by Story and Task (Target Categories Are Outside Parentheses, Foil Categories Are Inside Parentheses)

	Means of ranks		Ranks of mean ranks	
Story	Unitary (Basic level)	Disjunctive (Superordinate)	Unitary (Basic level)	Disjunctive (Superordinate)
		Rank by willingness to bet		
Yossi	1.6 (4.9)	3.6 (4.8)	1 (6)	3 (5)
Doron	2.3 (5.2)	3.9 (3.4)	1 (7)	3 (2)
Dorit	2.2 (4.8)	3.8 (4.3)	1 (7)	2 (4)
Alon	2.5 (5.1)	4.1 (4.8)	1 (7)	6 (4)
Asher	2.8 (4.7)	3.8 (3.2)	1 (6)	4 (2)
		Rank by probability		
Yossi	1.8 (4.4)	3.5 (4.7)	1 (4)	3 (5)
Doron	2.3 (5.0)	3.9 (3.2)	1 (6)	3 (2)
Dorit	2.5 (4.3)	4.0 (4.4)	1 (5)	3 (6)
Alon	3.0 (5.5)	4.8 (4.0)	1 (7)	6 (4)
Asher	3.0 (4.3)	3.5 (3.6)	1 (5)	2 (3)
		Rank by inclination to predict		
Yossi	1.8 (4.7)	3.2 (4.7)	1 (6)	3 (5)
Doron	2.1 (5.3)	4.2 (3.0)	1 (7)	3 (2)
Dorit	2.0 (5.1)	4.2 (4.6)	1 (7)	3 (6)
Alon	2.9 (4.5)	4.8 (4.7)	1 (5)	7 (6)
Asher	3.2 (4.4)	3.6 (4.2)	1 (6)	2 (5)
		Rank by suitability		
Yossi	1.6 (4.8)	3.2 (4.6)	1 (5)	3 (4)
Doron	2.7 (4.8)	3.7 (2.8)	1 (6)	3 (2)
Dorit	2.0 (5.0)	4.0 (4.2)	1 (7)	3 (4)
Alon	2.7 (5.4)	4.1 (4.4)	1 (7)	3 (6)
Asher	3.3 (4.8)	3.4 (4.0)	1 (7)	2 (4)

tive expectation, the overall disjunction fallacy rate (.77) was higher than the overall disjunction effect rate (.72), although the difference is not significant (z = 1.57). A similar picture emerges with respect to the foil categories: Here too the effect is, if anything, higher where it is fallacious (.36, probability and betting) than where it is not (.35, predicting and suitability) and is at its lowest for suitability (.32). Table 5 shows no evidence for any normative effects of task.

In contrast to the absence of a task effect, there does seem to be a story effect. Some of our stories (e.g., Yossi) simply "worked better" (i.e., elicited higher disjunction fallacy rates) than others (e.g., Asher). Insofar as this is a reflection of our uneven success in designing appropriate stimuli, this effect is of little interest. It becomes more interesting if we interpret it in light of the representativeness hypothesis. This hypothesis says that people judge probability by similarity, suitability, or representativeness. If so, the similarity between the results we obtained under the different sets of instructions simply reflects the fact that all subjects were essentially doing the same thing: Regardless of the formal criterion they were asked to rank the options by, they were all ranking the options by representativeness. We computed the correlation between the disjunction effect rates for the suitability task and the combined disjunction fallacy rates (i.e., between the rates reported on line 4 and 5 of Table 5). This correlation was r = .97.

The same pattern of results that emerges from the rates reported in Table 5 is repeated in the original ranks that subjects gave the critical options, reported in Table 6. We conducted seven separate 5 (Stories) × 4 (Tasks) analyses of variance (ANOVAs) using the following data: (a) the ranks of the basic level, or unitary, target category (leftmost column of Table 6); (b) the ranks of the superordinate target category (third column of Table 6); (c) the ranks of the basic level, or unitary, foil category (second column, in parentheses, of Table 6); (d) the ranks of the superordinate foil category (fourth column, in parentheses, of Table 6); (e) the difference of the ranks of the target unitary and superordinate categories; (f) the difference of the ranks of the foil unitary and superordinate categories; and (g) the difference of the previous two differences. In addition, we conducted a 5 (Stories) × 4 (Tasks) × 2 (Target, Foil) repeated measures ANOVA, in which category type (target vs. foil) was a repeated measure.

The ANOVA results are shown in Table 7. In all of them we found a highly significant main effect of the stories, but no main effect of the task, and no interaction. The absence of an effect for task confirms that all subjects were rendering similar judgments, in spite of being instructed to rank by different criteria. The significant effect found for the stories merely reflects, as we noted earlier, their uneven effectiveness. Happily, this unevenness did not interact with the tasks.

Additional evidence for the representativeness hypothesis can be found in Table 8. It summarizes an analysis that is based on the rankings of all seven options in our lists, not just the critical four (target pair + foil pair) reported in Table 7. For each of our five stories, we correlated the mean rankings of all seven options as elicited by the suitability criterion with the mean rankings elicited by each of the other criteria. The correlations are shown in Table 8. The median correlation is .91, and the combined correlation across all stories is in excess of .96. In comparison, Tversky and Kahneman (1983) found a correlation of .98 between mean rank by resemblance and mean rank by probability for Linda.

To be sure, correlations between mean ranks should be treated with circumspection, and even correlations as high as those reported here are not incompatible with large variability of individual rankings. However, in conjunction with the ANOVAs, and with the analyses carried out on the rates reported in Table 5, they support the interpretation that options were ranked much the same by all four criteria.

The similarities between the suitability rankings and the probability or the willingness-to-bet rankings are understandable in light of the representativeness hypothesis. The similarity between the probability rankings and the willingness-to-bet rankings is understandable in light of their normative equivalence. The similarity between the inclination-to-predict rankings and the suitability rankings requires no special explanation, because, unlike for probabilities, there is no contrast in the logic underlying the two. On the contrary, it is quite easy to defend the notion that the best prediction is not the one most likely to be true, but the one that offers the best combination of likelihood and informativeness.

Study 3

Method

In Study 1, subjects who ranked the options by their willingness to bet on them committed the disjunction fallacy at somewhat lower rates than those who ranked them by their probability. This minor, but systematic, effect disappeared in Study 2. Study 3 repeated the rank by willingness-to-bet task, but this time the bet was for real. Although we did not expect the move from betting for hypothetical money to betting for real money to have much impact, one clear advantage of betting for real money is that subjects are given a genuine incentive to suit themselves, not the experimenter, to do what they think is best for their interests rather than to try to guess what the experiment is about. Even if they believe that the experiment is about stereotypes, for example, or that they are really being asked for what is most reasonable to say rather than for what is most probably true, they can nonetheless exploit the options to their own benefit by ranking superordinate categories higher than their subsets. It may not make the experimenter happy, but it gives them a better shot at the reward.

The instructions were similar to those given in Studies 1 and 2, except for explicitly addressing themselves to the conditions under which the bet could be won, as follows:

> After you ranked all the options, we will check the option you ranked number 1 (namely, your first choice for betting). If you were right, you will immediately be awarded 100NS. If not, you will neither win, nor lose, anything. Unfortunately, we do not have the means to offer this bet to all of you, but at the end, we shall draw several people (according to the class size) who will actually bet for real. Each one of you has an equal chance of being drawn.

The questions in this study were administered in a class setting. The participants in this study were 25 seniors in philosophy and 65 seniors in economics. Each received two questions. After answering the questions, 2 subjects in the first class and 3 in the second were drawn by lottery to bet for real money (at the time 100NS was equivalent to about

Table 7
Analyses of Variance Results

Target	F_{task} ($dfs = 1, 3$)	p	F_{story} ($dfs = 1, 4$)	p	$F_{interaction}$ ($dfs = 1, 12$)	p
1. Unitary target	0.56	.64	13.18*	.0001	0.54	.89
2. Superordinate target	1.37	.25	7.65*	.0001	0.61	.83
3. Unitary foil	1.00	.39	2.82*	.02	1.09	.36
4. Superordinate foil	0.98	.40	20.49*	.0001	1.02	.43
5. Difference—target	1.05	.37	7.02*	.0001	0.67	.78
6. Difference—foil	0.95	.41	14.48*	.0001	1.52	.11
7. Difference of differences	0.26	.85	9.46*	.0001	1.66	.07
8. Difference—target & foil 3-dimensional	1.85	.14	11.00*	.0001	0.53	.89

* $p < .05$.

$45, a large sum considering that it was more than 10 times the hourly rate paid to students at the time, for a task that took only about 10 min).

In another group of 36 subjects (about equally divided between social science and education majors), a monetary reward was promised to every subject whose bet turned out correct, although for obvious budgetary reasons, this group was rewarded only 10NS each.

For the 100NS group, we chose from Study 2 the two strongest stimuli (in terms of eliciting disjunction fallacies), Yossi and Doron. Clearly, these are fictional characters, therefore we were at liberty to determine the "true" answer arbitrarily. One desirable feature of the disjunction rule is that its advantageousness cannot be undermined by any arbitrariness. Nonetheless, and although the settling of the bet was subsequent to the ranking of the options, we wanted the true answer to appear to subjects credible, not arbitrary. We did not want, however, to reward those subjects who committed the disjunction fallacy—namely, those who ranked the basic level target category number 1. Hence, we designated as the true answers "mathematical economics" for Yossi and "squash" for Doron. Although these answers were not directly listed among the options, anyone who chose to bet on the target superordinate category could win the bet. To check the suitability of mathematical economics and squash independently, we also ran a dual version of Yossi and Doron in which the original basic level target categories of "tennis" and "statistics" were replaced by "squash" and "mathematical economics," and "tennis" and "statistics" were designated as the true answers, respectively.

The 10NS group responded only to the original version of Gidi, who was the most effective of the Study 1 stimuli.

Table 8
Pearson Product-Moment Correlations Between Mean Ranking of the Seven Options by Suitability and by Other Criteria

Story	Suitability & willingness-to-bet	Suitability & probability	Suitability & tendency-to-predict
Yossi	.97*	.98*	.98*
Doron	.91*	.83*	.86*
Dorit	.91*	.96*	.94*
Alon	.92*	.87*	.68
Asher	.76*	.85*	.90*
Combined	.99*	.96*	.97*

* $p < .05$.

Results and Discussion

We checked informally that subjects who bet on the listed basic level target category did not feel unfairly tricked when told the true answer and that all respondents readily acknowledged that although the true answer was not in itself one of the seven listed options, the superordinate target category afforded a win of the bet.[6]

We gave Gidi's description to 36 subjects, of whom 72% committed the disjunction fallacy on the target category pair. The disjunction fallacy rate in Study 1's hypothetical betting condition was 74%—a small and nonsignificant difference ($z = .19$).

We gave the original versions of Yossi and Doron (as reported in Table 4) to 37 subjects, and 41 received the altered versions. For the original versions of both Yossi and Doron, the rates of the disjunction fallacy found in Study 3 were .68. For the altered versions, the rates were .81 for Yossi and .76 for Doron. Averaged across both versions, Yossi elicited the disjunction fallacy at a rate of .74, and Doron at a rate of .72. These rates are lower by less than 15% than those reported for these stories in the betting task in Study 2. This effect is not significant ($z = 1.59$ for Yossi, $z = 1.03$ for Doron). Moreover, because we selected the two most extreme stimuli from Study 2, some regression to the mean was to be expected.

Could the drop in the disjunction fallacy rate, small as it is, be due to subjects who recognize the normative appeal of the disjunction rule in this task? If so, we should expect to see dependence, perhaps even total overlap, between the subjects who gave normatively correct rankings in the two problems. On the other hand, if subjects are not systematically applying any insight learned in one question consistently to the other question, but rather answering both questions independently, then we would expect to find only 10% (32% × 32%) of the 37 original version subjects and 5% (19% × 24%) of the 41 altered version subjects answering both questions correctly—or about 6 subjects in our sample of 78.

In fact, however, we observed 8 such subjects—7 with the

[6] In one class, when we announced the true answer, one subject was heard to groan aloud: "Boy did I blow it! I chose statistics, when I could have chosen social science."

original version (i.e., 19%) and a single subject with the altered version (i.e., 2.4%). This is hardly higher than would be expected by independence (z = .98), and fails to support the possibility that subjects are applying the disjunction rule consistently. Moreover, if we add the foil pair, then the number of subjects who consistently obeyed the disjunction rule drops even farther (to 13% in the original version, and to 0% in the altered version—a total of 4 subjects).

General Discussion

The extension rule is perhaps the simplest and most transparent rule of probability theory, one whose validity even untrained and unsophisticated people accept and endorse. Hence, its violation is one of the most startling and dramatic errors of probabilistic reasoning. In Tversky and Kahneman's (1983) classic study, a single—and powerful—device was used to induce the fallacy: A representative feature was conjoined to an unrepresentative one. This addition increased the judged representativeness of the conjunction over that of its unrepresentative component, but the probability of the conjunction could, of course, only reduce the probability of each component. Thus, the modal judgment that P(A & B) is greater than P(A) provided striking and highly diagnostic evidence for the representativeness hypothesis.

The initial motivation behind the present study was to create a disjunction fallacy. It is well-known that, when categorizing things under conditions of certainty, people prefer basic level categories to their superordinate categories. If categorization under conditions of uncertainty were to exhibit the same preference, then disjunction fallacies might ensue. This would not only be an extension of the conjunction fallacy, but also would rely on an altogether different cognitive device than that evoked by representativeness.

To that end, the option lists we gave our subjects included a representative basic level category as well as its superordinate. We labeled this nested pair the *target pair.* To test whether the disjunction fallacy, if exhibited, really results from a general tendency to categorize at a basic level, the list also included an unrepresentative basic level category along with its superordinate. We labeled this nested pair the *foil pair.*

The two nested pairs elicited markedly different rankings. Whereas for the target pair the majority of subjects ranked the narrower category more likely than its superordinate, for the foil pair it was the reverse.[7] This pattern rules out an overall preference for basic level categories, but it could be compatible with the representativeness hypothesis. The rationale of representativeness does not require that conjunction fallacies be the only form of extension fallacies. Whenever an ordering of events by representativeness differs from their ordering by set inclusion, there is a potential for an extension fallacy to occur. To explain our results by representativeness, we need only show that the target basic level category was more representative than its superordinate, whereas the foil basic level category was less representative than its superordinate.

Rather than rely on our own a priori judgments of representativeness, in Study 2 we collected suitability rankings alongside probability and willingness-to-bet rankings for all the stories we used. Although some previous studies elicited the occasional representativeness judgment (e.g., Kahneman & Tversky, 1973; Tversky & Kahneman, 1983), this study correlated representativeness judgments with probability judgments systematically. Tables 5 and 6 show that the probability judgments, as well as the normatively equivalent willingness-to-bet judgments, closely followed the pattern exhibited by the suitability judgments, in conformity with the representativeness hypothesis. In particular, irrespective of the type of judgment rendered, the target basic level category was ranked above its superordinate for the target pair, and below it for the foil pair, just as required for a representativeness-based account of our results. Moreover, the disjunction fallacy rates are not merely correlated with the suitability disjunction effect rates, they are also quite similar to them.

Indeed, given the interaction we found between the disjunction fallacy rates and the type of nested pair, it seems not to matter whether the superordinate category really is disjunctive, or even whether it is at the Roschian superordinate level. Which of two events—even nested events—will seem more probable is better predicted by their representativeness than by their scope, or by the level in the category hierarchy in which they are located. Colleagues often ask us questions such as: What would have happened had you included subordinate categories in your option lists (e.g., "clay court tennis" for Doron, "Bayesian statistics" for Yossi)? What if the basic level category had looked *un*suitable, while its superordinate looked suitable (e.g., replacing South Africa with Algeria for Dorit)? What if one option had overlapped with another, or if there had been two subsets of a single superordinate? The representativeness hypothesis suggests a general answer for such speculative questions: If you want to predict how such manipulations would affect probability judgments, find out how they would affect representativeness judgments.

In addition to the probability and willingness-to-bet criteria, this study asked some subjects to rank the options by the degree to which they would be inclined to predict them. This criterion occupies an intermediary place between "how alike is it" and "how likely is it." It seems unobjectionable to predict, say, that Dorit wrote her letter from South Africa, because the description sounds more like South Africa than like any of the other options. After all, although there is another option in the list that is more likely to be true (i.e., Africa), maximizing the likelihood of being true is not what prediction is solely about. Prediction certainly relates more closely to maximizing the probability of being true than suitability does. In short, the mental distance between inclination to predict and either probability or suitability seems shorter than the mental distance between these two themselves.

Moreover, Yaniv and Foster's (1990) results suggest that estimators may not properly appreciate the tradeoff between informativeness of predictions and probability of being correct. In other words, they may believe that their predictions, although narrow and hence informative, are also highly likely to be cor-

[7] To be sure, *any* judgment that a category is more probable than its superordinate is a disjunction fallacy, even if only a minority of subjects exhibit it.

rect. This is an example of the kind of overconfidence often reported in the calibration literature. When making numerical estimates, people who are asked to give 95% confidence intervals (i.e., intervals that they are 95% sure will include the true value) are often right only about 50% of the time. They produce overly narrow estimated intervals with little awareness of the cost in justified confidence (for a survey see Lichtenstein, Fischhoff, & Phillips, 1982).

Recall Gavanski and Roskos-Ewoldsen's (1991) earlier mentioned claim that conjunction fallacies "stem primarily from the incorrect rules people use to combine probabilities." They also conceded, however, that "people almost certainly do sometimes judge conjunctions directly, through representativeness or other means" (p. 190). Our stimuli are a case in point, because they clearly could not be judged by combination rules. On the other hand, representativeness cannot account for all conjunction fallacies—for example, those found by Carlson and Yates (1989) or by Gavanski and Roskos-Ewoldsen. "Judgments of probability vary in the degree to which they follow a decompositional or a holistic approach and in the degree to which the assessment and the aggregation of probabilities are analytic or intuitive" (Tversky & Kahneman, 1983, p. 310). A pluralistic approach seems as much in order for extension fallacies as for other judgments of probability.

References

Adams, E. W. (1975). *The logic of conditionals.* Dordrecht, The Netherlands: D. Reidel.

Biela, A. (1986). Psychological patterns in predicting disjunction and conjunction of clinical symptoms. *Acta Psycologica, 61,* 183–195.

Carlson, B. W., & Yates, J. F. (1989). Disjunction errors in qualitative likelihood judgment. *Organizational Behavior and Human Decision Processes, 44,* 368–379.

Gavanski, I., & Roskos-Ewoldsen, D. R. (1991). Representativeness and conjoint probability. *Journal of Personality and Social Psychology, 61,* 181–194.

Grice, H. P. (1975). Logic and conversation. In G. Harman & D. Davidson (Eds.), *The logic of grammar.* Encino, CA: Dickinson.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3,* 430–454.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80,* 237–251.

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: The state of the art. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306–334). Cambridge; England: Cambridge University Press.

Morier, D. M., & Borgida, E. (1984). The conjunction fallacy: A task specific phenomenon? *Personality and Social Psychology Bulletin, 10,* 243–252.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Floyd (Eds.), *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Erlbaum.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8,* 382–439.

Shafir, E. B., Smith, E. E., & Osherson, D. N. (1990). Typicality and reasoning fallacies. *Memory and Cognition, 18,* 229–239.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84,* 327–352.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90,* 293–315.

Wells, G. L. (1985). The conjunction error and the representativeness heuristic. *Social Cognition, 3,* 266–279.

Wolford, G., Taylor, H. A., & Beck, J. R. (1990). The conjunction fallacy? *Memory and Cognition, 18,* 47–53.

Yaniv, I., & Foster, D. (1990). *On graininess of judgment: An accuracy-informativeness tradeoff.* Working paper, University of Chicago.

Received August 7, 1991
Revision received December 3, 1992
Accepted December 3, 1992 ■