THE CONJUNCTION EFFECT
AND CLINICAL JUDGMENT

HOWARD N. GARB
Wilford Hall Medical Center

Judgments made by clinicians may sometimes violate an axiom of probability theory, which dictates that neither the probability of event A, \( P(A) \) nor the probability of event B, \( P(B) \), can be less than the probability of the conjunction of events A and B, \( P(A \cap B) \). In the following study, 40 psychologists and psychology interns estimated the probability of one of their clients obtaining (a) a particular result on the Personality Assessment Inventory (PAI; Morey, 1991), (b) a particular result on the Minnesota Multiphasic Personality Inventory–2 (MMPI–2; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989), and (c) the same specified results on both the PAI and MMPI–2. Violations of the conjunction rule occurred for 58% of the clinicians and were relatively large in size.

A large literature in social and cognitive psychology is devoted to describing judgment processes in everyday life (e.g., Gigerenzer & Selten, 2002; Gilovich, Griffin, & Kahneman, 2002; Hastie & Dawes, 2001; Hogarth, 2001). Much of this literature describes cognitive heuristics and biases. *Heuristics* are simple rules that are descriptive of how judgments are made. *Cognitive biases* describe types of errors that people sometimes make.

The cognitive heuristics and biases that are descriptive of judgments made in everyday life may also be descriptive of the judgment processes of mental health professionals. By better understanding the judgment processes of mental health professionals, we can hope to improve clinical judgment and clinical practice. The purpose of the present study is to learn if a particular cognitive bias described by social and cognitive psychologists, the conjunction effect, occurs in clinical practice.
The conjunction effect was described by Tversky and Kahneman (1983). If a person believes that the probability of two events occurring, \( P(A \cap B) \), is greater than the probability of just one of those events occurring (e.g., \( P(A) \)) then the conjunction effect is said to have occurred. According to probability theory, neither the probability of event A, \( P(A) \), nor the probability of event B, \( P(B) \), can be less than the probability of the conjunction of events A and B, \( P(A \cap B) \). For example, if one flips a coin twice, the likelihood of obtaining “heads” on the first coin toss, \( P(A) \), cannot be less than the probability of obtaining “heads” on the first coin toss and “tails” on the second \( P(A \cap B) \).

To study the conjunction effect, Tversky and Kahneman (1983) had participants make ratings for several vignettes including one that described a social activist named Linda:

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice and also participated in antinuclear demonstrations. Please rank the following statements by their probability, using 1 for the most probable and 8 for the least probable:

1. Linda is a teacher in an elementary school.
2. Linda works in a bookstore and takes yoga classes.
3. Linda is active in the feminist movement.
4. Linda is a psychiatric social worker.
5. Linda is a member of the League of Women Voters.
6. Linda is a bank teller.
7. Linda is an insurance salesperson.
8. Linda is a bank teller and is active in the feminist movement.

Eighty–five percent of the participants rated alternative h as more probable than f, in violation of probability theory.

Findings on the conjunction effect are not without controversy (Gavanski & Roskos-Ewoldsen, 1991; Gigerenzer, 1996; Kahneman & Tversky, 1996). A number of investigators in the area of cognitive psychology have argued that the conjunction effect may occur in some instances because of the ambiguity of language (Shafir & LeBoeuf, 2002). That is, respondents may understand the judgment task differently than the experimenters. For example, when instructed to make ratings, some participants may infer from the statement “Linda is a bank teller” that “Linda is a bank teller who is not a feminist.” Tversky and Kahneman (1983) anticipated this criticism, and they conducted an additional experiment that included the alternative: “Linda is a bank teller whether or not she is active in the feminist movement.” When this alternative was
included, the magnitude of the conjunction fallacy was diminished but
the fallacy persisted among a majority of participants. Other investiga-
tors have also investigated whether the conjunction effect is due to the
ambiguity of language. In general, conjunction effects have frequently,
but not always, been found (e.g., Dulaney & Hilton, 1991; Mellers,
Hertwig, & Kahneman, 2001; Morier & Borgida, 1984; Sides, Osherson,
Bonini, & Viale, 2002; Stolarz–Fantino, Fantino, Zizzo, & Wen, 2003;
Tentori, Bonini, & Osherson, 2004).

The conjunction effect has received, and continues to receive, a great
deal of attention in social and cognitive psychology, in part because its
occurrence suggests that judgments sometimes violate a central tenet of
probability theory (Shafir & LeBoeuf, 2002). In contrast, although more
than 1,000 studies have been conducted on clinical judgment in the area
of personality assessment, most of them have been on the reliability and
validity of judgments, and none have been on the conjunction effect
(Garb, 1998, 2005; Garb, Lilienfeld, & Fowler, 2005). The goal of this
study is to learn if the conjunction effect occurs when mental health pro-
fessionals make judgments. By having clinicians predict a client’s test re-
results, one can avoid some of the problems that have been encountered
with studying everyday judgments. With the Linda problem, respond-
ents may have been making inferences at an unconscious or automatic
level about whether it is likely that Linda is, or is not, a feminist even
when they were simply told that Linda is a bank teller. This type of un-
conscious processing should not occur in the present study. For the task
used in the present study, one can easily assume that a client can be
given one test but not another. It seems unlikely that respondents will
make assumptions about a client’s results on a test that was not
administered.

In the following study, clinicians will estimate the likelihood of a client
receiving a particular set of test results. When examining the test results
for a client, clinicians often form opinions about how likely the results
seem. In addition, in describing how case conferences could be made
more meaningful, Meehl (1973) encouraged psychologists to make pre-
dictions regarding test results, and then afterward discuss why some re-
results may not have turned out as predicted:

The important point is that the inferences arrived at in the staff confer-
ence would include predictions about what the . . . MMPI said, and these
agreements or discrepancies should constitute differential reinforce-
ments for adequate versus inadequate clinical behavior by the partici-
pants. I may, of course, still think that I am right and that the MMPI is
wrong; but it is a fact that I mispredicted . . . the personality profile based
upon the MMPI. (p. 295)
More recently, Silver (2001) has also encouraged psychologists to predict a client’s test results as a first step in test interpretation.

In this study, a trade-off is made between the clinical relevance of the judgment task and being able to obtain relatively unambiguous results. Judging whether test results seem to describe a client is not among the most important tasks that clinicians perform. However, other judgment tasks were not chosen because they would not allow for results that are as unambiguous. For example, the conjunction effect may occur when one instructs psychologists to estimate the probability of a client having: (a) a dysthymic disorder, (b) an antisocial personality disorder, and (c) both a dysthymic disorder and an antisocial personality disorder. However, when instructed to make ratings, some participants may infer from the statement “The client has a dysthymic disorder” that “The client has a dysthymic disorder and not an antisocial personality disorder.” Thus, when asked to rate the likelihood of dysthymic disorder, the psychologist may actually be rating the likelihood of dysthymic disorder and “antisocial personality disorder not present.” The likelihood of this conjunction can be greater than the likelihood of the conjunction of dysthymic disorder and antisocial personality disorder, without violating the axioms of probability theory.

METHOD

JUDGES

U.S. Air Force psychologists and psychology residents (interns) were asked to participate in this study. This sample was chosen for two reasons. First, this group is thought to have strong clinical skills. In fact, the psychology department at Wilford Hall Medical Center (WHMC), with one of the largest internship programs in the nation, recently received the Outstanding Training Program Award from the Association for the Advancement of Behavior Therapy (now the Association for Behavioral and Cognitive Therapies). Second, there is an emphasis on cooperation and support within the Air Force, and it was expected that a high response rate would be obtained with this sample. Psychologists and psychology residents who were at WHMC in 2003 or 2004 or who had been at one of four additional bases in 2004 were asked to participate in this study.

Participation was voluntary. Participants remained anonymous: identifying information was not recorded. No financial incentive to participate was offered. Participants were instructed to list their professional status (psychology resident or staff psychologist), gender, and type of degree program (Ph.D. or Psy.D.). Additional demographic data were
not recorded because it was important to preserve the anonymity of the participants.

PROCEDURE

Each clinician was instructed to think of a client whom he or she knows well and has seen for evaluation or treatment. He or she was told that all of their ratings were to be made for this client. Having clinicians rate one of their own clients ensures that ratings will be made for a wide range of individuals. This should make it easier to generalize from the results of the study. That is, findings will not be due to the use of an idiosyncratic vignette. Clinicians were to make a series of probability (likelihood) ratings. They were told that probability (likelihood) ratings range from 0 to 1. They were also told that a rating of 0 should indicate a zero-percent chance of an event occurring, a rating of .5 should indicate a 50–50 chance, and a rating of 1 should indicate that the event is certain to occur.

Each clinician was to assume that his or her client was given the Personality Assessment Inventory (PAI; Morey, 1991). If the clinician was unfamiliar with this test, he or she was to assume that it is a generally valid measure. The clinician was to estimate the likelihood of his or her client obtaining the following results on the PAI:

(a) The test results suggest that the client is combative with other people. ____
(b) The test results suggest that the client is likely to be impulsive and emotionally labile. ____
(c) The test results suggest that the client tends to be angry and suspicious while at the same time being anxious and needy. ____

Next, each clinician was to assume that his or her client was not given the PAI, but instead was given the Minnesota Multiphasic Personality Inventory–2 (MMPI–2; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989). Each clinician was to estimate the likelihood of the following results on the MMPI–2:

(d) The client is experiencing a moderate to severe level of emotional distress characterized by dysphoria, guilt, and anxiety. The client is likely to overreact to minor stress. The client lacks self-confidence and feels insecure, inadequate, and inferior. ____

Finally, each clinician was to assume that his or her client was given both personality inventories. The clinician was then instructed to estimate the likelihood of the following results:
Upon finishing, each clinician was encouraged to review all of his or her ratings and change them if he or she wished. The conjunction effect will have occurred if: the likelihood rating for e is greater than the likelihood rating for a or d, the likelihood rating for f is greater than the likelihood rating for b or d, or the likelihood rating for g is greater than the likelihood rating for c or d. Thus, for each clinician, the conjunction effect could occur as often as six times.

RESULTS

All of the psychologists and psychology residents at five U.S. Air Force bases were asked to participate, except for one who was deployed and another three who participated in a pilot study and gave feedback on an earlier version of the questionnaire. Of the 56 psychologists and psychology residents, 40 completed questionnaires for a 71% response rate. This is an excellent response rate compared to response rates obtained in other clinical judgment studies (e.g., Ford & Widiger, 1989). Twelve of the clinicians were psychology residents and 28 were staff psychologists. Nineteen were female, and 19 were male (two did not indicate gender). Finally, 27 had been in a Ph.D. program, and 11 had been in a Psy.D. program (two did not indicate program).

The conjunction effect occurred for 58% (23 of 40) of the clinicians. As mentioned earlier, the conjunction effect could occur zero to six times for
any individual clinician. On average, it occurred 1.3 times for every clinician (standard deviation = 1.6), and 2.2 times for every clinician for whom at least one conjunction effect occurred (standard deviation = 1.5). The shape of the distribution of conjunction effects deviated significantly from the normal distribution (Kolmogorov–Smirnov z(N = 40) = 1.71, p < .01), precluding the use of parametric significance testing. The occurrence of conjunction effects was not related to professional status (staff vs. resident), χ²(2, N = 40) = .63, p = > .05, gender, χ²(2, N = 40) = .96, p = > .05, or type of graduate program (Ph.D. vs. Psy.D.), χ²(2, N = 40) = 1.48, p = > .05.

Conjunction effects varied in size. To describe the magnitude of a conjunction effect, one can subtract the probability of a conjunct, P(A), from the probability of the conjunction, P(A ∩ B). For example, if the probability of a PAI interpretation was estimated to be .20 and the probability estimate of the same PAI interpretation coupled with the MMPI–2 interpretation was estimated to be .45, then the magnitude of the conjunction effect would be .25. A positive value indicates that the conjunction effect has occurred. In the present study, when a conjunction effect occurred, the average magnitude of that effect was .26 (standard deviation = .21). The distribution of the conjunction effects is presented in Table 1. (For 240 comparisons, conjunction effects occurred 51 times. It is these 51 conjunction effects that are described in Table 1.)

**DISCUSSION**

The results of this study suggest that some psychologists violate an axiom of probability theory when they make judgments. The conjunction

<table>
<thead>
<tr>
<th>Magnitude</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>.05</td>
<td>3</td>
</tr>
<tr>
<td>.10</td>
<td>19</td>
</tr>
<tr>
<td>.20</td>
<td>12</td>
</tr>
<tr>
<td>.30</td>
<td>2</td>
</tr>
<tr>
<td>.40</td>
<td>1</td>
</tr>
<tr>
<td>.50</td>
<td>11</td>
</tr>
<tr>
<td>.70</td>
<td>1</td>
</tr>
<tr>
<td>.75</td>
<td>1</td>
</tr>
<tr>
<td>1.0</td>
<td>1</td>
</tr>
</tbody>
</table>
effect occurred for over half of the psychologists and psychology residents, and the average magnitude of these effects was relatively large. Results did not differ as a function of professional status, gender, or type of graduate program.

Strengths and limitations of the study can be described. First, although the response rate is high, the sample is only modest in size. With a larger sample size and additional measures, it might be possible to identify differences among the clinicians that are related to the validity of their judgments. Second, this study used a unique sample: U. S. Air Force psychologists and residents. It is possible that the results from this study will not generalize to other psychologists, although this seems unlikely in part because Air Force psychology personnel receive training at civilian graduate schools across the country. Third, with regard to a strength of the study, every clinician made ratings for a different client: one whom they presumably knew well. This should help to bolster external validity. Finally, although the task used in the present study is important (Meehl, 1973, p. 295), one can wonder if conjunction effects occur for additional clinical tasks. A reason that predicting test results was used in the present study is because it allows us to collect relatively unambiguous results: it is not difficult for clinicians to imagine a client being given one test but not another. In contrast, for the Linda vignette (Tversky & Kahneman, 1983), it may have been difficult for participants to think of Linda as being a bank clerk without also drawing unconscious inferences that she was likely to not be a feminist. Given that relatively unambiguous results were obtained for the conjunction effect for one clinical judgment task, it seems likely that conjunction effects may also occur for other tasks. That is, if some clinicians violate the axioms of probability theory when making ratings for one judgment task, they may also violate those same axioms when making ratings for other tasks.

In conclusion, when interpreting the test results for a client, clinicians will want to avoid violating the axioms of probability theory. In particular, they should be careful to ensure that conjunction effects do not occur. Graduate students should be taught about the conjunction effect, not because we know that its occurrence is widespread across a range of tasks in clinical practice, but because we know that it can occur on occasion. Learning about the conjunction effect should help to sharpen critical thinking.
REFERENCES


