Oldham, G.R., & Fried, Y. (1987). Employee reactions to workspace characteristics. *Journal of Applied Psychology*, 72, 75-80.

Park, O.S., Sims, H.P., & Motowidlo, S.J. (1986). Affect in organizations. In H. Sims & D. Gioia (Eds.), *The thinking organization* (pp. 1-44). San Francisco: Jossey Bass.

Sinclair, R.C. (1988). Mood, categorization, breadth, and performance appraisal: The effects of order of information acquisition and affective state on halo, accuracy, information retrieval, and evaluations. *Organizational Behavior and Human Decision Processes*, 42, 22-46.

Thomas, K.W. (1976). Conflict and conflict management. In M. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (pp. 889-935). Chicago: Rand McNally.

Thomas, K.W. (1977). Toward multi-dimensional values in teaching: The example of conflict behavior. *Academy of Management Review*, 2, 484-490.

Thomas, K.W. (in press) Conflict and negotiation processes in organizations. In M.D. Dunnette (Ed.), *Handbook of industrial and organizational psychology* (2nd ed.). Chicago: Rand McNally.

Van Toller, S., & Dodd, G.H. (1988). *Perfumery: The psychology and biology of fragrance*. London: Chapman Hall.

Williams, K.J., & Keating, C.W. (1987). *Affect and the processing of performance information*. Paper presented at the meetings of the Society of Industrial and Organizational Psychology, Atlanta.

Zillmann, D., Baron, R.A., & Tamborini, R. (1981). The social costs of smoking: Effects of tobacco smoke on hostile behavior. *Journal of Applied Social Psychology*, 11, 548-561.

The Psychology of Contraceptive Surprises: Cumulative Risk and Contraceptive Effectiveness[1]

HARRIET SHAKLEE AND BARUCH FISCHHOFF[2]

Eugene Research Institute

Two studies investigated young adults' expectations about long-term contraceptive effectiveness. Subjects were told about five hypothetical methods of contraception varying in reported effectiveness, which was expressed in terms of the likelihood of avoiding pregnancy for base periods of 1 year (Experiment 1), 5 years, or 10 years (Experiment 2) of use. For each method, subjects estimated the likelihood that a woman would avoid pregnancy while using it for periods ranging from 1 month to 15 years, and then rated how satisfied they would be with it. For nearly half of the subjects, estimates of cumulative effectiveness did not decline as time period increased. Those subjects who did realize that cumulative effectiveness declined over time estimated rates that declined too slowly for methods of modest and low reliability, and at rates that were too similar for methods differing in effectiveness. Subjects were overly optimistic about effectiveness for time periods longer than the base period, and overly pessimistic about effectiveness for shorter time periods. Not surprisingly given these results, subjects expressed more satisfaction when a method's effectiveness was expressed in shorter base periods. Such faulty understanding of the long-term implications of contraceptive effectiveness information may undermine people's abilities to make informed contraceptive choices

Birth control may be thought of as a continuing process of risk management. Sexual intercourse carries a risk of conception, which can be reduced through the use of various contraceptive methods, differing in how effectively they control fertility. The effectiveness of contraceptive methods can be known to some degree of precision through scientific studies. However, when people make decisions about contraception, they must rely on their own subjective understanding of effectiveness.

Various studies have shown that beliefs about effectiveness influence which method of contraception, if any, people choose (e.g., Houser and Beckman, 1978; Miller, 1986; Tanfer & Rosenbaum, 1986). Unfortunately, those beliefs are often quite inaccurate for teenagers, adults, and even contraception counselors (e.g., Americans exaggerate, 1985; Tanfer &

[1]Partial support for this research was provided by National Science Foundation grants BNS-8509673 and SES-8715564 awarded to the first and second authors, respectively, and by a grant from the Carnegie Council on Adolescent Development to the second author. We are grateful to Brad Bell for his assistance in data collection and to Lita Furby for her comments on earlier drafts of this paper, and to Rosa Lio for manuscript preparation.

[2]Requests for reprints should be sent to Baruch Fischhoff, who is now at the Department of Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, PA 15213.

Journal of Applied Social Psychology, 1990, 20, 5, pp. 385-403

Rosenbaum, 1986; Trussell, Faden,& Hatcher, 1976). Such misconceptions may lead people to choices that fail to reflect their own true priorities. The present studies investigate how people understand one particular aspect of contraceptive effectiveness: how risks accumulate over repeated exposure.

Many risks in life have a relatively low probability for any single exposure. However, these small probabilities add up over repeated exposures to create a substantial overall risk. For example, the risk of having an accident on any individual car trip is so low that the effort of using a seat belt may seem unjustified. However, over a lifetime of rides, the chances of being in at least one accident are considerable. Indeed, Slovic, Fischhoff, and Lichtenstein (1978; also Schwalm & Slovic, 1982) found that people considered seat belt use to be more attractive when the risks of driving were described from a long-term perspective. The effectiveness of this manipulation suggests that people fail to understand the long-term implications of information about short-term driving risk.

The risk of conception, too, is relatively low on a single occasion of sexual intercourse, but accumulates over repeated occasions. Without contraception, the cumulative risk of becoming pregnant increases quite rapidly with each new exposure. However, even with many commonly available methods of contraception, the long-term risk of at least one unplanned pregnancy is quite high. For example, for a birth control method with a one-year reliability of .96, one would expect 4 out of 100 women to become pregnant during one year of use. However, if the same 100 women continued to use the method, each additional year would bring more pregnancies. The long-term likelihood of avoiding pregnancy over multiple years of contraceptive use is equal to the product of the likelihood of avoiding pregnancy in each year of use. If we assume that the effectiveness of a method is constant across years of use, that same contraceptive would be expected to produce at least one conception among 19% of couples using it over a 5-year interval, 34% over a 10-year period, and 46% over 15 years of use.

Clearly, it is critical to understand how risk accumulates over time. Unfortunately, studies in other content domains have found limits in intuitive reasoning processes that would lead to undue optimism about the long-term protection provided by contraceptive use. Avoiding conception is like a compound gamble, which one "wins" only by having a favorable outcome on each of several plays. However, studies show that people generally overestimate their chances of winning such gambles (Bar-Hillel, 1973; Cohen, Chesnick, & Haran, 1971; Slovic, 1969), as a result of underestimating the rate at which the likelihood of winning declines over multiple plays. People have enough faith in these perceptions that they may prefer a compound gamble to a simple gamble that has an objectively greater likelihood of winning. Related research by Svenson (1984, 1985) demonstrates people's limited ability to assess the cumulative risk associated with exposure to compound environmental hazards.

In the context of contraception, a comparable bias would mean that people who learn short-term effectiveness rates would tend to overestimate the likelihood of avoiding pregnancy over longer time periods. Effectiveness information is typically presented for one year of use: that is, the proportion of women using the method for a year who successfully avoid pregnancy. In fact, though, the time period is often omitted, suggesting perhaps to recipients that it is irrelevant to the likelihood of experiencing a contraceptive failure.

Experiment 1 investigated young adults' expectations about the long-term reliability of contraceptive methods when offered information about annual reliability. Subjects were told about hypothetical contraceptive methods, with effectiveness values in the range of those associated with the commonly available methods of birth control. For half of the subjects, those values were described as pertaining to one year of contraceptive use; the remaining subjects received the same values, but without a specified time period. All subjects were asked to estimate the cumulative effectiveness of those methods for periods varying from 1 month to 15 years of use, then to rate their relative satisfaction with each method.

Experiment 1

Method

Procedure. Subjects were introduced to the concept of contraceptive reliability, defined as the proportion of women who would avoid becoming pregnant when using a given method of birth control. They were then told about five hypothetical methods of contraception, with reliabilities of .99, .96, .93, .90, and .87. For subjects in the *one-year condition*, these reliability values were described as pertaining to one year of method use. The *unspecified condition* said nothing about the time period.

For each method of birth control, subjects estimated the proportion of women who would avoid becoming pregnant after using the method for 1 month, 1 year, 5 years, 10 years, and 15 years. A sample item for the 1-year condition (deleting "for 1 year" produces the comparable item in the unspecified form):

(A) The reliability of method A is .99 for 1 year. That is, if 100 women used this form of birth control for 1 year, 99 of these women would not become pregnant. Please estimate the probability that a woman who consistently used this method as directed would successfully avoid becoming pregnant over a:

1 month period _____; 1 year period _____;
5 year period _____; 10 year period _____;
15 year period _____.

After rating all methods for their cumulative effectiveness, subjects then evaluated each method:

If you were interested in birth control, how satisfied would you be with the reliability of this method?

0	1	2	3	4	5	6	7	8	9	10
not at all satisfied										very satisfied

Would you consider this method to be effective enough to use for birth control? Yes No (circle one)

Because these evaluation questions followed the judgments of cumulative risk, the process of making those judgments may have affected subjects' satisfaction ratings. To examine this possibility, some subjects in the unspecified and one-year conditions received the evaluation questions alone.

Subjects. Subjects were 76 females and 40 males, ranging from 18 to 27 years of age, who responded to an advertisement in a university newspaper offering payment for participation in studies on judgment and decision making. The subjects were divided equally between the one-year and unspecified conditions. Within the one-year and unspecified conditions, 21 and 19 subjects, respectively, made evaluation judgments alone, with the remaining subjects making both cumulative effectiveness and evaluation judgments.

Results

Effectiveness. Each subject's pattern of judgments of cumulative effectiveness was classified according to whether it (a) decreased, (b) increased, (c) was flat, or (d) was nonmonotonic as the time period increased. Subjects were so classified if their judgment functions had the same form for at least four of the five contraceptive methods, a criterion met by nearly all of the subjects (95%). Among them, only 53% of subjects in the one-year condition and 59% in the unspecified condition (appropriately) estimated lower probabilities for increased time periods. Flat cumulative effectiveness functions were the next most common pattern (28% and 24% of classifiable subjects in the one-year and unspecified conditions, respectively), followed by increasing (17% and 11%) and nonmonotonic functions (5% and 3%). Distributions of these classifications were similar for male and female subjects.

Judgments were analyzed further for those subjects whose cumulative effectiveness judgments declined with time period. Each subject's judgments were analyzed in terms of the slope and intercept of the cumulative effectiveness function over the five time periods for each of the five methods. Differences in slopes would indicate different expectations regarding the rate of decline over time. Different intercepts would indicate differential expectations about short-term risk. These values were entered into a MANOVA, with condition (unspecified, one-year) and subject gender as between-subject factors, contraceptive method as a within-subject factor, and slope and intercept as the two dependent variables. These analyses showed that subjects' judgments did differ as a function of method reliability, multivariate $F(8, 24) = 28.71$, $p < .001$, with significant effects for both the slope, $F(4, 28) = 22.14$, $p < .001$, and the intercept, $F(4, 28) = 44.01$, $p < .001$. The interaction between gender and method approached significance in the multivariate analysis, $F(8, 24) = 2.11$, $p < .10$, but not in the univariate analyses for either the slope or the intercept. There were no significant differences associated with condition, meaning that subjects treated unspecified effectiveness information in the same way as they treated the same values when identified as pertaining to one year of use.

Figure 1 combines mean judgments of the unspecified and one-year groups, for subjects with monotonically decreasing responses. The significant intercept effect is reflected in differing estimates in effectiveness for one month of use. Slope differences are reflected in the modest fanning of the curves. Both of these patterns are in the normatively appropriate direction, but far smaller than they should be. Figure 1b shows computed effectiveness estimates, assuming a constant annual rate of contraceptive effectiveness across years of use.

Comparing the two panels of Figure 1, one can see that subjects' function slopes are much too similar for methods of different reliability. This produced subjective functions that are too flat for methods of annual effectiveness of .96-.87, but too steep for the most effective method (.99).

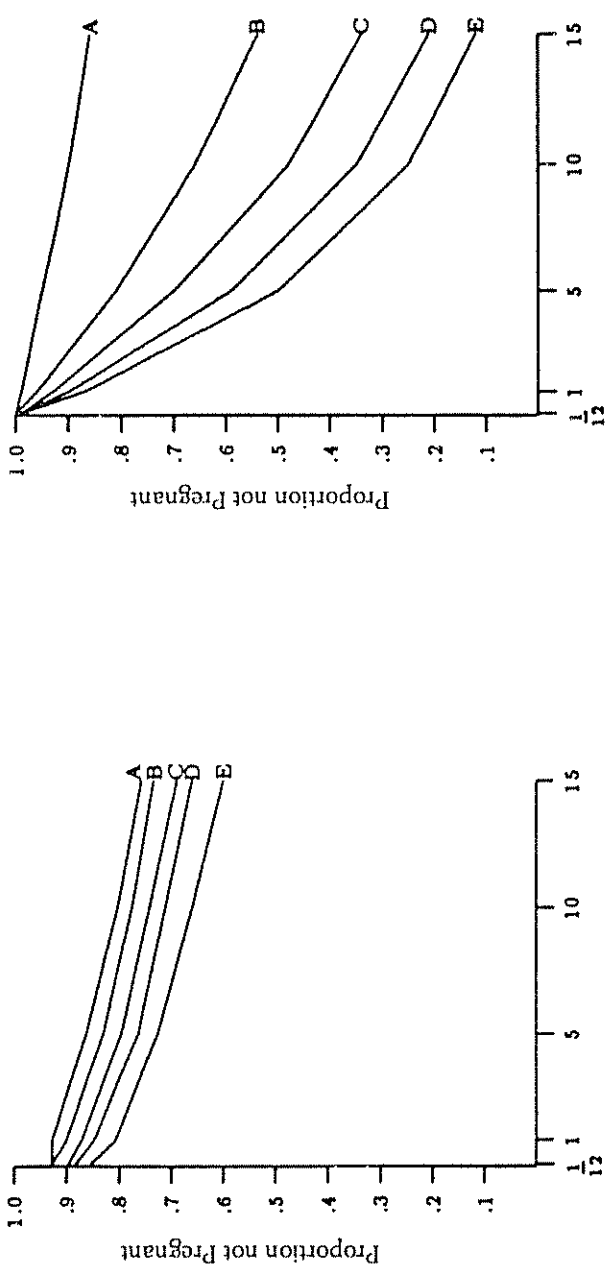

Figure 1. Proportion of women who would not become pregnant over time while using contraceptive methods at 5 effectiveness levels (A-E, .99 to .87): (A) mean estimate for subjects whose cumulative effectiveness functions declined over time; (B) computed value.

The near-significant interaction between gender and method reflected somewhat steeper and less parallel judgment functions for males than for females; thus, among this slim majority of subjects who knew that effectiveness declined with time, males did a slightly better job of estimating the actual difference in rates. Overall, however, even these more normative subjects were insufficiently sensitive to the effects of both time period and differences in annual reliability. Subjects both underestimated the long-term effectiveness of the most reliable method (.99) and greatly overestimated the long-term reliability of less reliable methods (.96-.87). The net effect was to treat methods of different reliability as too equivalent in long-term outcome.

Method evaluation. After making these cumulative effectiveness judgments, subjects indicated how satisfied they would be with a method having each level of reliability, and whether they considered it effective enough to use for birth control. A second set of subjects within each condition made these evaluative ratings alone. Responses to these two evaluation questions were entered into parallel ANOVAs, with between-subject factors of gender, judgment condition (one-year vs. unspecified), and whether cumulative effectiveness judgments had been made (yes or no), and with the different effectiveness levels as a within- subject factor. As one would expect, subjects' satisfaction ratings increased significantly with method effectiveness (3.86, 5.06, 6.37, 7.57, 9.02; $F(4, 105) = 124.63$, $p < .001$). More interestingly, subjects were significantly more satisfied with contraceptive methods in the unspecified than in the one-year form, $F(1, 108) = 4.48$, $p < .05$, even though subjects in the two conditions judged cumulative effectiveness similarly. However, a significant condition-by-gender interaction, $F(1, 108) = 5.42$, $p < .05$, shows that this effect was restricted to the male subjects, who were less satisfied when the time period was given as one year (5.40) than when it was left unspecified (6.91). For female subjects, there was no difference (6.64 vs. 6.56). Parallel to these satisfaction ratings, the percentage of subjects who considered the methods to be effective enough to use declined sharply with effectiveness (84.6, 27.8, 10.7, 2.1 and 2.6). Because of the overwhelming rejection of methods with reliabilities of .90 and .87, judgments for these methods were dropped from further analysis. Analyses of the remaining judgments showed a significant effect for effectiveness, $F(2, 109) = 75.02$, $p < .001$, but no other significant effects. Most notably, for both response formats, comparisons between conditions showed that the process of making cumulative effectiveness judgments had no effect on method evaluation.

Discussion

Previous studies of judgments about compound gambles suggested that people would have difficulty understanding the long-term implications of information about short-term risk. In particular, we expected people to underestimate the rate at which risk accumulates over time. The results of Experiment 1 clearly supported these expectations for judgments about long-term contraceptive outcomes. The greatest problems in understanding were evidenced by the 41%-47% of the subjects who did not realize that cumulative effectiveness declines consistently with increasing time. However, even subjects who understand this fact were generally too optimistic about long-term contraceptive effectiveness. Optimism was especially strong for methods of modest or poor reliability. Subjects somewhat underestimated the long-term effectiveness of the most reliable method.

As shown in Figure 1a, cumulative effectiveness judgments were too similar for methods differing in reliability, rather than fanning out as shown in Figure 1b. The greatest discrepancies are with the longest time periods, with subjects failing to realize the extent to which relatively small differences in short-term effectiveness mount up to very large differences over time. Adding the responses of the large group of subjects whose cumulative effectiveness judgments did not decline over time would produce even greater discrepancies between judged and calculated values.

Also interesting is the similarity in subjects' judgments for the unspecified and one-year forms. The only statistically significant effect of specifying the base period was that males (but not females) gave lower satisfaction ratings to the contraceptive methods in the one-year form. Females may be more likely than males to know that reported effectiveness rates typically pertain to one year of use, hence to make that inference in the absence of a specified time period. Alternatively, males may better appreciate the importance of time period for long-term outcomes, hence better realize the relevance of base-period information. Such sensitivity may also have been seen in the marginally significant tendency for males to produce steeper and less parallel cumulative effectiveness judgment curves.

Thus far, we have evaluated subjects' judgments in terms of cumulative risk estimates computed under the assumption that annual contraceptive effectiveness is constant across years of use. The information typically offered to potential contraceptive users would give them no reason to think otherwise. However, research has found that the annual

effectiveness of contraceptive methods tends to increase in the second year of use (Vaughn, Trussell, Menken, & Jones, 1977) and may improve further thereafter. Might our subjects' relatively flat judgment functions for cumulative risk reflect their knowledge of this fact?

This account seems unlikely for several reasons. One is that scientific information about year-to-year changes in contraceptive effectiveness is not widely available. A second reason is that this interpretation cannot account for subjects' tendency to underestimate the long-term effectiveness of the most reliable method (.99). Finally, nearly half of our subjects did not realize that cumulative effectiveness declines consistently over time, an expectation that clearly conflicts with reality.

Experiment 2

Results of Experiment 1 suggest that a short-term perspective on risk may promote unrealistic optimism about long-term outcomes. Perhaps people would appreciate the problem of cumulative risk better if they were introduced to risks from a long-term perspective. Experiment 2 investigated this possibility.

Method

Procedure. Subjects were told about five different hypothetical methods of contraception with annual effectiveness levels comparable to those in Experiment 1. However, in this case, contraceptive effectiveness was expressed in terms of either *5 years* or *10 years* of method use. For example, the method with a one-year reliability of .99 was represented as having an effectiveness level of .95 for 5 years ($= .99^5$) and .90 for a period of 10 years ($= .99^{10}$); the method with a one-year reliability of .87 was represented as having a 5-year reliability of .50 ($= .87^5$) and a 10-year effectiveness of .25 ($= .87^{10}$). As in Experiment 1, subjects estimated the proportion of women who would avoid getting pregnant for periods of 1 month, 1 year, 5 years, 10 years, and 15 years. Subjects also evaluated each contraceptive method, using the two questions from Experiment 1. Again, some subjects in each time period condition evaluated the methods without first making judgments of cumulative risk.

Subjects. Subjects were 63 females and 52 males who were recruited by the same methods as in Experiment 1. Fifty-six subjects participated in the 5-year condition and 58 subjects in the 10-year condition, with 21 and 22 subjects in the two conditions, respectively, making only evaluation judgments.

Result

Effectiveness. As in Experiment 1, subjects were classified according to whether their cumulative effectiveness judgments decreased, increased, were flat, or nonmonotonic as time period increased. In the 5-year condition, 91.4% of the subjects showed a single pattern for at least four of the five methods; in the 10- year condition 83.3% were so classifiable. Among these subjects, cumulative effectiveness judgments declined over time for 65.6% of the 5-year and 50.0% of the 10-year subjects. Flat judgment functions were the next most frequent pattern (16.7% and 33.3% of the 5- and 10-year subjects, respectively). Increasing functions where shown by 9.4% and 10.0% of the 5- and 10-year subjects, respectively, and nonmonotonic functions by 9.4% and 6.7% of the 5- and 10-year subjects, respectively. This distribution of response patterns was similar for males and females, and parallels that of Experiment 1.

Subjects with monotonically decreasing cumulative effectiveness judgments were included in further analyses. Because both subjects and methods in the present study are comparable to those in the one-year condition in Experiment 1, we incorporated their data into the present analyses. Figure 2 shows subjects' cumulative effectiveness judgment functions in each of the base-period conditions.[3] As in Experiment 1, we analyzed these judgments in terms of the slopes and intercepts of cumulative effectiveness functions for each of the five methods in a MANOVA including between-subject factors of base-period condition (1-, 5-, and 10-year) and gender, and the within-subject factor of method reliability (.99-.87). Significant effects for base-period condition would mean that subjects judge cumulative risks differently depending on how they are initially expressed. Differences in the judgments of short-term risk would emerge in significant intercept effects, and differences in judging the rate of risk accumulation would be evident as slope effects.

As in Experiment 1, these analyses showed an effect of method reliability, multivariate $F(8, 36) = 160.55$, $p < .001$, with significant effects in both the slope and intercept, $F(4, 40) = 87.15$ and 130.27, respectively, $p < .001$ for both.[4] Thus, overall, this subset estimated different

[3]The slight initial upward tends in Figure 2 in cumulative effectiveness functions for the 99 method in the 1- and 10-year conditions may seem anomalous among subjects who have been preselected for their monotonically decreasing functions. However, subjects were included if they had decreasing functions for four out of the five methods, opening a possibility that group functions may be nonmonotonic, as evident here. In each condition, the upward trend is due to a single subject's low estimate (< 10) for the 1- month period.

[4]This is not an independent replication of the method-reliability effect from Experiment 1, because the present analyses include data from that study.

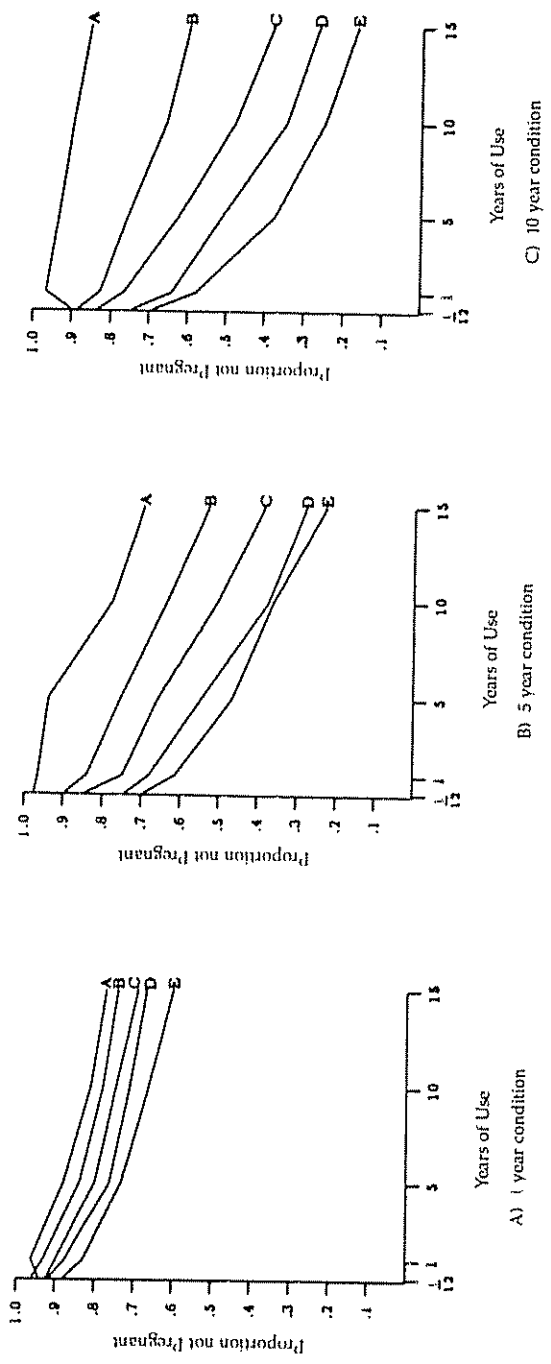

Figure 2. Mean estimated proportion of women who would not become pregnant over time while using contraceptive methods at 5 effectiveness levels (A-E, annual effectiveness of .99 to .87), including only those subjects whose cumulative effectiveness functions declined over time. Effectiveness provided to subjects in terms of (A) 1 year, (B) 5 years, (C) 10 years.

one-month risks and different rates of risk accumulation for these different methods. Judgment functions were also influenced by judgment condition, multivariate $F(4, 86) = 4.92$, $p < .001$, with a significant effect in the intercept alone, $F(2, 43) = 9.73$, $p < .001$. As can be seen in Figure 2, there was a greater spread in one-month reliability estimates for the 5- and 10-year conditions, in which subjects received initial rates that were spread more widely. These latter groups also showed steeper judgment functions for most methods. Both these main effects are qualified by an interaction between condition and method effectiveness, multivariate $F(16, 74) = 6.14$, $p < .001$, with effects significant in both slope and intercept, $F(8, 80) = 9.44$ and 9.84, respectively, $p < .001$, and by a three-way interaction between condition, method, and gender, multivariate $F(16, 74) = 1.82$, $p < .05$, with effects significant for both the slope and intercept, $F(8, 80) = 2.74$ and 2.51, respectively, $p < .05$. The interaction between gender and method was also significant, multivariate $F(8, 36) = 2.89$, $p < .05$, significant for both the slope and intercept, $F(4, 40) = 4.99$ and 5.50, respectively, $p < .01$.

The pattern of these effects on function intercepts showed that differences in estimates of short-term effectiveness among the five contraceptive methods were greater for the 5- and 10-year base conditions than for the 1-year condition, with this effect of base period stronger for female than for male subjects. In addition, for females, but not males, effects of the base-period condition on judgments of short-term risk were greatest for the less effective methods, when the shift in base-period would result in the greatest differences between conditions in the risk value presented to the subjects. As mentioned, subjects' judgment functions were generally steeper in the 5- and 10-year conditions than in the 1-year condition. Within each time period condition, females' judgment functions were flatter than those of males for all except the .99 method. Females' judgment functions for the five contraceptive methods were also more parallel than those of males. Comparisons between these judgment functions and the calculated functions (Figure 1B) show that males' and females' judgments of cumulative effectiveness are too parallel for the five methods and too flat for all except the most reliable method. The gender differences (which are not shown in the figures) involved greater deviations for the female subjects

Method evaluation. The two evaluation ratings for each subject were analyzed separately in ANOVAs that included between-subject factors of gender, base-period condition (1-, 5-, and 10-year), and whether cumulative effectiveness judgments had been made, as well as the within-subject factor of method reliability. Subjects' ratings of method satisfaction

Table 1

Satisfaction Ratings for Contraceptive Methods as a Function of Annual Effectiveness and Base Period

Base period	Annual effectiveness				
	0.99	0.96	0.93	0.90	0.87
One year (Exp. 1)	8.69	7.27	5.86	4.77	3.49
Five years (Exp. 2)	8.39	5.80	4.36	1.92	1.55
Ten years (Exp. 2)	7.29	3.63	2.21	1.21	0.71

Note. 0 = not at all satisfied; 10 = very satisfied.

showed main effects of method effectiveness, $F(4, 158) = 229.96$, $p < .001$, and judgment condition, $F(2, 161) = 38.38$, $p < .001$, and a significant interaction between the two factors, $F(8, 316) = 11.63$, $p < .001$. Mean judgments appear in Table 1. As before, satisfaction ratings decreased with method reliability, with this effect proving stronger for the 5- and 10-year base periods that provided information on the cumulative risk that subjects in the one-year base period were unable to generate for themselves. Subjects also preferred each contraceptive method when its reliability was expressed in the higher value associated with shorter time periods. This effect was weakest for the most reliable method (one-year effectiveness of .99), where varying the base period of risk expression produces the least differences in the effectiveness value presented to subjects. Gender and judgment condition also showed a significant interaction, $F(2, 161) = 3.42$, $p < .05$, with female satisfaction ratings more influenced by base period of risk expression (means over all five methods = 6.63, 4.11, 2.38, for the 1-, 5, and 10-year conditions, respectively) than were male judgments (5.40, 4.70, 2.89).

As in Experiment 1, subjects overwhelmingly rejected the two least reliable methods when judging whether they were effective enough to use (see Table 2). Analyses of judgments for the remaining three methods show significant effects of base-period condition and method, $F(2, 162) = 10.16$; $F(2, 161) = 169.50$, respectively, $p < .001$, and an interaction between the two factors, $F(4, 322) = 6.01$, $p < .001$. These reflected a greater willingness to accept the .99 and .96 methods when their reliabilities were

Table 2

Percentage of Subjects Who Rate Methods as Effective Enough to Use as a Function of Annual Effectiveness and Base Period

Base Period	Annual effectiveness				
	.99	.96	.93	0.90	0.87
One year (Exp. 1)	88.0	38.1	10.5	1.7	0.0
Five years (Exp. 2)	91.5	16.0	2.5	0.0	0.0
Ten years (Exp. 2)	71.1	6.9	10.3	0.0	1.7

expressed in terms of shorter base periods. As in Experiment 1, neither measure of satisfaction was influenced by whether subjects had previously rated the methods for cumulative effectiveness.

Discussion

In Experiment 1, subjects told the short-term effectiveness of contraceptive methods overestimated their long-term effectiveness. Their judgment functions were generally too flat, underestimating the rate of risk accumulation, and too parallel, treating methods varying in effectiveness as too equivalent in long-term outcome. Experiment 2 found that introducing risk from a long-term perspective still produced cumulative risk judgment functions that were relatively flat and parallel. In both the 5- and 10-year conditions, subjects underestimated risks for time periods longer than the base period and overestimated risks for shorter time periods. Estimates of short-term effectiveness were generally lower when effectiveness was expressed in terms of longer time periods, especially for methods of poorer reliability. Thus, although a short-term perspective on risk promoted unrealistic optimism about long-term outcomes, a long-term perspective may have produced undue pessimism about short-term risk. As in Experiment 1, judgment problems were greatest for the sizeable groups of subjects (34.4%-50.0%) who failed to realize that cumulative risk declines consistently over time (and were excluded from other statistical analyses).

With all three base conditions, methods varying substantially in reliability produced nearly parallel curves for judgments of cumulative effectiveness rather than the fan indicated in Figure 1B. One general account of this pattern is in terms of the anchoring-and-adjustment heuristic (Tversky & Kahneman, 1974). Individuals using this judgment heuristic take a salient value as an anchor which they then adjust, typically by an insufficient amount, in order to reach their judgments. The effectiveness value provided to the present subjects is an obvious anchor, from which subjects adjust up (for shorter periods) or down (for longer periods). Inadequate adjustment would produce overly flat curves. The near-parallel nature of these curves would mean that the size of those adjustments is largely independent of the level of annual risk.

Evaluations of the contraceptive methods paralleled subjects' estimates of cumulative effectiveness. Both when rating their satisfaction and when judging whether the methods were effective enough to use, subjects preferred methods when effectiveness levels were expressed in shorter base periods (which produced higher estimates of cumulative effectiveness). For satisfaction ratings, this effect was strongest for the less reliable methods, for which the base period manipulation had the greatest effect on expected cumulative risk.

More clearly evident in Experiment 2 are sex differences in these judgments. Although males were no more likely than females to realize that cumulative effectiveness declines over time, among those subjects who did produce decreasing functions, males' judgments of cumulative effectiveness were steeper and less parallel than those of females, a pattern more closely approximating the calculated functions. Males' cumulative effectiveness judgments were also less affected by the (formally irrelevant) base period of risk expression, a gender difference which was strongest for the least reliable methods of birth control. Parallel gender effects are seen in subjects' evaluations of the contraceptive methods, with base period of risk expression showing less effect on the satisfaction ratings of males than of females. These differences come despite the greater substantive familiarity with contraceptive methods that females might be expected to have. If robust, a somewhat greater ability to compute cumulative probabilities intuitively might be another reflection of the slight superiority in mathematical abilities that males have shown on standardized tests (e.g., Deaux, 1985; Maccoby & Jacklin, 1974).

As in Experiment 1, the process of making cumulative risk judgments had no effect on people's satisfaction ratings for these methods. Possibly, people think about cumulative risk (albeit inaccurately) even without an explicit prompt to do so. Clearly, a willingness to think about long-term risk has limited value without the ability to do so.

The specifics of these results must depend on the particular stimuli used here, in terms of cover story (contraceptives), base period, reliabilities, and estimated periods. Nonetheless, the basic pattern is one that has been observed repeatedly in related tasks. As mentioned in the introduction, people have been found to overestimate their chances of winning compound gambles, where, as here, they cannot tolerate a single failure (Bar-Hillel, 1973; Cohen et al., 1971; Slovic, 1969). People also underestimate rates of change for other nonlinear processes, including projecting the future of exponentially growing series (e.g., Wagenaar & Sagaria, 1975) and the long-term effects of inflation on the value of money (Bates & Gabor, 1986; Kemp, 1987). The generality of these patterns suggests a deep-rooted problem which will appear in various guises and be hard to eliminate in intuitive judgments.

Cumulative risk and contraceptive choice. These results may shed some light on recent trends in contraceptive use in the United States. Our subjects overestimated the long-term effectiveness of methods reported to have modest-to-poor reliability, but underestimated the long-term effectiveness of the most reliable method. Among currently available birth control options, oral contraceptives offer the highest reliability, often rated as high as .99 (annually). By contrast, methods such as the diaphragm, condom, and foam are less effective, with reliabilities generally below .90. Studies have found that people who change contraceptives generally shift to a method with lower reliability than the one they had been using (Grady, Hayward, Billy, & Florey, in press). For example, there has been a recent shift among young married couples from oral contraceptives to barrier methods of contraception (Pratt, Mosher, Bachrach, & Horn, 1984; Tanfer and Horn, 1985), apparently reflecting concern over the health hazards associated with the pill.[5] However, the present results suggest that such couples will underestimate the sacrifice in long-term cumulative effectiveness that they are making by shifting to the less reliable methods. A better understanding of these cumulative risks of an unplanned pregnancy might enhance the attractiveness of the pill (or other highly effective methods) to such couples. No research evidence can tell couples how to weigh health concerns versus effectiveness when choosing a contraceptive. However, couples need to understand all risks involved if they are to make decisions in their own best interest. Results of the present studies show systematic biases in how people translate risks from one time period to another. As a result, the responsible strategy for risk communicators is to present effectiveness estimates for all time periods relevant to people's decision.[6, 7]

[5]The accuracy of these perceptions would be a suitable topic for another study (Fortney, 1988)

These results also suggest that people may be overly surprised by the contraceptive failures that they do experience. That is, people with excessive confidence in their contraceptive method are less likely to attribute unplanned pregnancies to its incomplete reliability. In their search for some other causes, the members of a couple may come to blame one another's care in contraceptive practice, undermining their relationship at a time when it is already stressed by the unplanned pregnancy.

The present evidence of bias in expectations regarding long-term contraceptive effectiveness suggests a basic problem in achieving informed consent for contraceptive choices. Houser and Beckman (1978), Miller (1986), Tanfer and Rosenbaum (1986) and others have found that effectiveness is an important consideration in people's decisions about contraception. Aiding those decisions is one motivation for research designed to estimate the effectiveness of currently available contraceptive methods (Trussell & Kost, in press). Making this information available does not by itself, however, allow individuals to make informed choices among alternative methods—if even technically correct information can be misinterpreted. Ensuring informed consent requires sensitivity to the psychology of people's decision making as well as to its informational content (Fischhoff, 1985). Otherwise, informed consent is more illusory than real.

References

Americans exaggerate health risks of the pill, underestimate effectiveness of all contraceptives. (1985). *Family Planning Perspectives, 17*, 128-129.

[6]In such communications, the relationships between risks over different exposure periods may seem surprising to recipients. One risk (suggested by an anonymous reviewer) is that if people cannot reconcile their intuitions, they may try to "beat the odds" associated with using any but the most reliable methods for many years by switching regularly in order to take advantage of the small risk for one year's use. If people do generate this strategy spontaneously, cautioning against it might be an additional responsibility of risk communicators.

[7]In another reflection of this problem, Linville, Fischhoff, and Fischer (1989) asked subjects to estimate the probability of the HIV virus being transmitted from a male to a female in one case and in 100 cases of protected sex. The mean one-case estimate was 10%, much higher than current scientific estimates (Fineberg, 1988). The mean 100 case estimate of 25% was much lower than it should have been (given these subjects' one-case estimate), suggesting an inability to estimate how rapidly small risks add up. It was, however, much closer to scientific estimates. A public health communication that presented this badly needed information from only one exposure perspective could badly mislead individuals making decisions for other exposure periods.

Bar-Hillel, M. (1973). On the subjective probability of compound events. *Organizational Behavior and Human Performance*, **9**, 396-406.

Bates, J., & Gabor, A. (1986). Price perception in creeping inflation: Report on an enquiry. *Journal of Economic Psychology*, **7**, 291-314.

Cohen, J., Chesnick, E.I., & Haran, D. (1971). The evaluation of compound probabilities in sequential order. *Nature*, **32**, 414- 416.

Deaux, K. (1985). Sex and gender. *Annual Review of Psychology*, **36**, 49-81.

Fineberg, H.V. (1988). Education to prevent AIDS: Prospects and obstacles. *Science*, **239**, 592-596.

Fischoff, B. (1985). Managing risk perceptions. *Issues in Science and Technology*, **2**, 88-96.

Fortney, J. (1988). Contraception—A lifelong perspective. In *Dying for Love*. Washington, DC: National Council for International Health.

Grady, W., Hayward, M., Billy, J., & Florey, F. (in press). Contraceptive switching among currently married women in the United States. *Journal of Biosocial Science*.

Houser, B., & Beckman, L. (1978). Examination of contraceptive perceptions and usage among Los Angeles County women. *Contraception*, **18**, 7-18.

Kemp, S. (19877). Estimation of past prices. *Journal of Economic Psychology*, **8**, 181-189.

Linville, P., Fischhoff, B., & Fischer, G. (1989). *Judging the risks of AIDS*. Manuscript submitted for publication.

Maccoby, E., & Jacklin, C. (1974). *The psychology of sex differences*. Stanford, CA: Stanford University Press.

Miller, W. (1986). Why some women fail to use their contraceptive method: A psychological investigation. *Family Planning Perspectives*, **18**, 27-32.

Pratt, W., Mosher, W., Bachrach, C., & Horn, M. (1984). Understanding U.S. fertility: Findings from the National Survey of Family Growth, Cycle III. *Population Bulletin*, **39**, 1-43.

Schwalm, N.D., & Slovic, P. (1982). *Development and test of a motivational approach and materials for increasing use of restraints*. PFTR-1100-82-3. Woodland Hills, CA: Perceptronics, Inc.

Slovic, P. (1969). Manipulating the attractiveness of a gamble without changing its expected value. *Journal of Experimental Psychology*, **79**, 139-145.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1978). Accident probabilities and seat belt usage: A psychological perspective. *Accident Analysis and Prevention*, **10**, 281-285.

Svenson, O. (1984). Cognitive processes in judging cumulative risk over different periods of time. *Organizational Behavior and Human Performance*, **33**, 22-41.

Svenson, O. (1985). Cognitive strategies in a complex judgment task: Analyses of concurrent verbal reports and judgments of cumulated risk over different exposure times. *Organizational Behavior and Human Decision Processes*, **36**, 1-5.

Tanfer, K., & Horn, M. (1985). Contraceptive use, pregnancy and fertility patterns among single American women in their 20's. *Family Planning Perspectives*, **17**, 10-19.

Tanfer, K., & Rosenbaum, E. (1986). Contraceptive perceptions and method choice among young single women in the U.S. *Studies in Family Planning*, **17**, 269-277.

Trussell, J., Faden, R., & Hatcher, R. (1976). Efficacy information in contraceptive counseling: Those little white lies. *American Journal of Public Health*, **66**, 761-767.

Trussell, J., & Kost, K. (in press). Contraceptive failure in the United States: A critical review of the literature. *Studies in Family Planning*.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, **185**, 1124-1131.

Vaughan, B., Trussell, J., Menken, J., & Jones, E. (1977). Contraceptive failure among married women in the United States, 1970-1978. *Family Planning Perspectives*, **9**, 251-258.

Wagenaar, W., & Sagaria, S. (1975). Misperception of exponential growth. *Perception & Psychophysics*, **18**, 416-422.