Journal of Experimental Psychology:
Human Perception and Performance
1978, Vol. 4, No. 2, 330-344

Fault Trees: Sensitivity of Estimated Failure Probabilities to Problem Representation

Baruch Fischhoff, Paul Slovic, and Sarah Lichtenstein
Decision Research, A Branch of Perceptronics
Eugene, Oregon

Fault trees represent problem situations by organizing "things that could go wrong" into functional categories. Such trees are essential devices for analyzing and evaluating the fallibility of complex systems. They follow many different formats, sometimes by design, other times inadvertently. The present study examined the effects of varying three aspects of fault tree structure on the evaluation of a fault tree for the event "a car fails to start." The fault trees studied had four to eight branches, including "battery charge insufficient," "fuel system defective," and "all other problems." Major results were as follows: (a) People were quite insensitive to what had been left out of a fault tree, (b) increasing the amount of detail for the tree as a whole or just for some of its branches produced small effects on perceptions, and (c) the perceived importance of a particular branch was increased by presenting it in pieces (i.e., as two separate component branches). Insensitivity to omissions was found with both college student subjects and experienced garage mechanics. Aside from their relevance for the study of problem solving, such results may have important implications for (a) how best to inform the public about technological risks and to involve it in policy decisions and (b) how experts should perform fault tree analyses of the risks from technological systems.

Many problems involve some form of troubleshooting: Something goes wrong, and the problem solver attempts to figure out why by first listing and then checking possible causes. The problem solver's representation of the problem (and possible causes) can be constructed from scratch, retrieved from memory, or adopted from an external source, perhaps with supplementary information from memory. A common representation for such problems is a "fault tree," which organizes possible sources of trouble into a branching structure. Figure 1 presents a fault tree for the event "a car fails to start long enough to delay the driver for 1 minute." The top row presents the problem, and the next row indicates the major systems whose failure might be relevant. Below each major system are listed specific contributing failures. Sources used in constructing this particular tree were *Petersen's Basic Auto Repair Manual, The Dodge Colt Service Manual, The Chevelle Owner's Guide,* and several experienced mechanics. Using analogous sources of ex-

---

We would like to thank Barbara Combs and Gerry Hanson for their help in conducting these experiments; David Collins for technical advice on designing fault trees; and Nancy Collins, Phyllis Kemper, and Peggy Roecker for drafting the many versions of the fault tree that we used.

This research was supported by the Advanced Research Projects Agency of the Department of Defense and was monitored by the Office of Naval Research under Contract N00014-76-C-0074 to Oregon Research Institute (ARPA Order No. 2449) from Decisions and Designs, Inc., and to Perceptronics (ARPA Order No. 3052) from Decisions and Designs, Inc. (Subcontract 76-030-0714).

Requests for reprints should be sent to Baruch Fischhoff, Decision Research, A Branch of Perceptronics, Inc., 1201 Oak Street, Eugene, Oregon 97401.

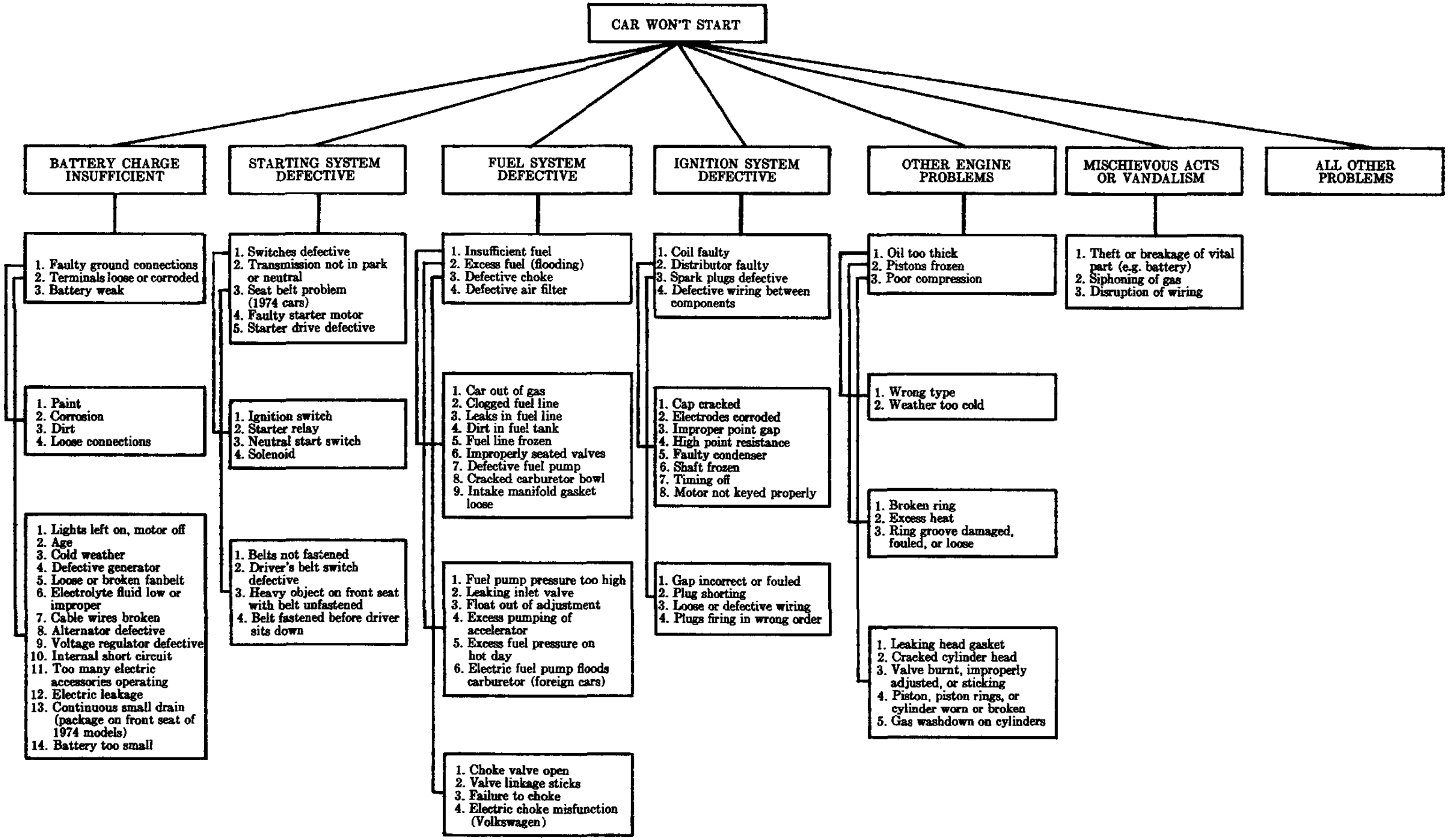

Figure 1. A possible fault tree for discovering why a car won't start.

pertise, one could construct a tree headed by "3-month old infant cries more than 5 minutes," "bank statement does not match checkbook," "star pitcher fails to report for spring training," "SAC bomber recall system fails," or "nuclear power reactor core melts."

Once constructed, the tree serves as a guide for the problem solver who might ask: Which system is more likely to be the source of trouble? What information would allow me to check out the fault most quickly? How well do I know this system (i.e., how authoritative is the fault tree), and should I be tinkering with it myself? Fault trees are used not only in analyzing systems that have gone astray but also in attempting to design fail-safe systems. Knowing how things go wrong is a prerequisite to drafting directives on what to do right. Fault trees have been instrumental in the design of technological systems from spaceships to nuclear power plants (Atomic Energy Commission, 1975; Green & Bourne, 1972; Bryan, Note 1).

Fault trees are also used to estimate failure rates for complex systems when historical data for the system as a whole are unavailable. Probabilities are assigned to each of the pathways to failure and are then combined to provide an overall failure rate. Such analyses were the primary methodological tools in the $3 million Rasmussen study, which assessed the probability of a catastrophic loss-of-coolant accident in a nuclear power reactor (Atomic Energy Commission, 1975). Fault tree analysis has, however, come under attack from critics who question whether it is methodologically sound enough to be used as a basis for decisions of great consequence (e.g., Bryan, Note 1). One major concern of the critics is that omission of relevant pathways due to ignorance, poor memory, or lack of imagination would lead to an underestimation of failure probabilities (Kendall, 1975).

As with many other kinds of problems (Newell & Simon, 1972), any troubleshooting situation can be represented in a variety of ways. For example, in constructing the fault tree in Figure 1, it was necessary to decide how much detail to provide for each system; whether to present minor systems like "mischievous acts" separately or to lump them with "all other problems"; whether the four items grouped in "fuel system defective" actually belong together or whether the last two might best be listed separately with a heading like "carburetion problems"; whether to use the graphic display or an outline; and whether to use this level of specificity, or more, or less. Making these decisions will depend on considerations like the purpose of the analysis, the amount of knowledge available about the issue in question, and how much of an effect each of these aspects of presentation has on the way people evaluate the system.

The studies reported here look at the impact of three such "arbitrary" aspects of the way in which fault trees are presented and the way in which they are evaluated. These are: (a) what is listed specifically and what is left to all other problems, (b) how much detail is presented for the various "branches" (systems) of the fault tree, and (c) how various systems are grouped into branches.

Aside from their interest for students of problem solving, such questions may be important for those concerned with the management of risks in our society. The lay public is increasingly called upon to decide whether the risks from various technological systems are acceptable in the light of the accompanying benefits. The risks from many proposed projects (e.g., nuclear power, liquid natural gas, recombinant DNA research) are varied and complex. Typically, they are presented to the public by technical experts who in one way or another use fault tree representation. In preparing their presentation, they must make arbitrary decisions like those listed above.

Understanding what difference these decisions make might help us understand (a) whether the public is being properly informed (e.g., Is risk information being presented in ways leading to its subjective overestimation or underestimation?), (b) how technical information can be communicated so that it is perceived most veridically, and

(c) what possibilities exist for manipulating perceptions through judicious fault tree presentation. If the perceptions of experts are also affected by variations in tree representation, we might learn something about how to improve the basic methodology of a technique used for momentous decisions.

Experiment 1: Pruning the Tree

However exhaustive one would like to be in the design of a fault tree, at some point it is necessary to stop listing alternatives and combine the remainder into a category labeled something like *all other problems*. In theory, listing every possible cause is an endless chore, and in practice, listing many implausible causes could overwhelm the fault tree's designers and viewers with bizarre possibilities that might divert attention from the primary possibilities. In Experiment 1, we studied the sensitivity of an evaluator to fault tree components that have been omitted. In this initial investigation, we looked for gross effects, those obtained by deleting substantial portions of a tree without specifically indicating their absence. In the real world of design and disaster, elements can be left out either inadvertently (the designer lacks the appropriate knowledge or imagination) or consciously (the designer collapses certain categories into the "other" category for simplicity; the designer omits certain pathways to make the system appear safer).

Two contrary hypotheses can be advanced for the impact of deleting major components of a fault tree on a viewer's judgment of its completeness.

One hypothesis suggests that when major items are deleted, it is quite likely that the absence of at least one item will be detected. Once such an omission has been uncovered, the entire analysis is discredited and the proportion left out is exaggerated. (Imagine your reaction to discovering that a purported fault tree for "car won't start" contains no mention of battery failure.) Many public discussions regarding nuclear power and other technological risks would seem to support this hypothesis. Members of the public scrutinize a fault tree prepared by a technical expert, discover (what seems to them to be) an important omission, and doubt the quality of the entire analysis and the competence of the analyst (e.g., Barrager, Judd, & North, 1976; Birnbaum, Wong, & Wong, 1976; Settle & Golden, 1974).

In contrast with this "credibility" hypothesis, the "availability" hypothesis (Tversky & Kahneman, 1973) suggests that what is out of sight is also out of mind. Problems that are not mentioned explicitly may not be thought of. People may not realize what is missing and, therefore, may overestimate the completeness of the analysis. Suspecting this to be the case, critics of technological projects who discover omissions in fault trees have cast doubt not on the competence but on the integrity of the analysts, charging the analysts with deliberately omitting problems to induce the public to underestimate the total risks associated with the project.

A third hypothesis is that people have such well-defined frequency representations in their minds (Howell, 1973) that they will appropriately realize what is missing. It is also possible that availability and credibility effects might cancel one another, producing appropriate judgments of all other problems.

Method

Stimulus. The basic stimulus for all studies reported here is the car won't start fault tree of Figure 1. It was constructed by consulting a variety of shop and repair manuals, mechanics, and car buffs to make it as complete as possible.[1] We distinguish between three levels of detail: *Level 1,* which presents just the system names (e.g., "battery charge insufficient"); *Level 2,* which lists three to five component problems for each system (e.g., (a) faulty ground connections; (b) terminal loose or corroded; (c) battery weak); and *Level 3,* which provides fuller detail for each of the component problems that could be elaborated, that is, everything shown in Figure 1.

---

[1] To our chagrin, the first starting failure encountered by anyone connected with this study was not included in our fault tree: an ignition key not turning because the steering wheel lock had caught the ignition switch.

Subjects in Experiment 1 were presented with Level 2 detail.

Design. Four separate groups of subjects participated. Two groups received the full, unpruned tree of Level 2. They differed only in that the experimenter read aloud the entire tree to one group ($n = 58$); the other group ($n = 35$) read the tree by themselves, without an enforced amount of time for examining it. Each of the two other groups was given a different pruned tree; one group ($n = 29$) saw a tree that was missing the starting, ignition, and mischief branches, and the other group ($n = 26$) saw a pruned tree that lacked the battery, fuel, and other engine branches. Since the first two groups showed no difference in response measures, the self-paced read-to-oneself administration was used for the two pruned tree groups.

Procedure. Before studying the fault tree, subjects were told:

> Every day, across the United States, millions of drivers perform the act of getting into an automobile, inserting a key in the ignition switch, and attempting to start the engine. Sometimes the engine fails to start, and the trip is delayed. We'd like you to think about the various problems that might be serious enough to cause a car to fail to start so that the driver's trip is delayed for at least 1 minute.
>
> The chart on the next page is intended to help you think about this problem. It shows six [three] major deficiencies that cause a car's engine to fail to start. These major categories probably don't cover all possibilities, so we've included a seventh [fourth] category, All Other Problems.
>
> Please examine this diagram carefully and answer the following question:
>
> For every 100 times that a trip is delayed due to "starting failure," estimate, on the average, how many of those delays are caused by each of the seven [four] factors. Make your estimates on the blank lines next to the factors named below. Your estimates should sum to 100.

The numbers above in brackets were given to subjects who saw the pruned trees.

After assigning proportion, all subjects were asked:

> Please answer the following question: Over the next 1,000 times in the U.S. that drivers attempt to start their cars, how many times will the drivers experience delays in starting the engine serious enough to delay their departure by at least 1 minute?
>
> Answer: ______ times out of 1,000

They were also asked a number of questions about the extent of their current and past experience with cars.

Subjects. One hundred forty-eight persons who responded to an advertisement in the University of Oregon student newspaper participated. They were assigned to an experimental group according to their preference for experiment date and hour.

Results

The difference in the mean proportion of problems assigned to each branch for the self-paced group and the read-aloud group was minimal. (Mean absolute difference across the seven branches was .018.) Results for these groups were combined.

If pruned tree subjects were sensitive to what had been omitted in the trees they studied, the proportion of problems that they attributed to "other" would equal the sum of the proportions of problems attributed to the pruned branches and to "other" by subjects who saw the full tree. If the availability hypothesis were correct, "other" would be assigned a lower proportion in the pruned tree; if the credibility hypothesis were correct, that proportion would be greater.

Table 1 presents the mean proportion of starting failures attributed to each branch and to "other." Pruned tree subjects clearly failed to appreciate what had been left out. For pruned tree Group 1, "other" should have increased by a factor of six (from .078 to .468) to reflect the proportion of failures due to starting, ignition, and mischief. Instead, "other" was only doubled, whereas the importance of the three systems that were mentioned was substantially increased. A similar picture emerged with the second pruned tree. Although these results clearly favor the availability over the credibility hypothesis, subjects did show some sensitivity to what had been omitted. The proportion of problems attributed to "other" was significantly greater for both pruned tree groups than for the unpruned tree group. (For both t tests, $p < .001$.) In addition, pruned tree Group 2, which saw a tree with more important branches deleted, assigned a higher proportion of problems to "other"

Table 1
Results from Experiment 1: Pruning the Tree

Group	*n*	Mean proportion of starting failures attributed to: Battery	Starting system	Fuel system	Ignition system	Engine	Mischief	Other	*Mdn* no. starting failures in 1,000
Unpruned tree	93	.264	.195	.193	.144	.076	.051	.078	75
Pruned tree 1	29	.432	—	.309	—	.116	—	.140[a]	59
Pruned tree 2	26	—	.357	—	.343	—	.073	.227[b]	76

Note. A dash indicates that the branch was deleted.
[a] Should be .468.
[b] Should be .611.

than did pruned tree Group 1, $t(53) = 2.59$.

Only 1 subject of the 55 in the two pruned tree groups assigned to "other" a proportion of problems greater than or equal to the sum of the proportions of problems assigned to the missing branches (plus "other") by the unpruned tree group. There was no tendency for subjects who rated themselves higher in expertise (on a 5-point scale) to assign a higher proportion of problems to "other" in any of the groups.

There were no significant differences between the pruned and unpruned tree groups in the median likelihood that a randomly selected car would not start. This is a bit surprising, because the support for the availability explanation noted above would suggest that mentioning more specific sources of difficulty to unpruned tree subjects would increase the availability of starting failure. Medians were used because of the presence of some extremely high estimates (e.g., 600 starting failures in 1,000 tries), which would have unduly affected the means. The relative frequency of such unrealistic estimates (even among experienced drivers) suggests that this measure was not entirely successful, making results obtained with it somewhat dubious.

Discussion

Pruned tree subjects clearly did not appreciate how much had been left out, and, as a consequence, they overestimated the exhaustiveness of the branches they saw. One could speculate that this effect would be even more pronounced with fault trees concerning technical systems less familiar than the present.

Although the enormity of the present effect suggests that it may be quite robust, one might wonder whether it was due, at least in part, to unpruned tree subjects either (a) assuming that the experimenters were competent and thus had provided a reasonably complete tree or (b) not attending sufficiently to the "other" branch. Experiment 2 was conducted to explore these possibilities.

Experiment 2: Focusing on All Other Problems

Method

Experiment 1 was repeated with several changes. To focus subjects' attention on what was left out, a paragraph was added to the instructions saying, "In particular, we'd like you to consider its completeness. That is, what proportion of the possible reasons for a car's not starting are left out, to be included in the category *all other problems?*" Their proportion estimation task was reduced to answering:

> For every 100 times that a trip is delayed due to "starting failure," estimate, on the average, how many of those delays are caused by factors not included in the chart. ______ out of every 100 cases would fall in the category *all other problems.*

Because of the difficulty that Experiment 1 subjects seemed to have had in assessing the likelihood of starting failure with a randomly selected car, the question was expanded by (a) asking them not only about a randomly selected car but also about their own car and (b) asking not only for an absolute judgment (how many times out of 1,000 attempts) but also for a relative judgment. (How many times more or less likely is a starting failure than a flat tire?)

Eighty-two subjects were recruited as in Experiment 1. The fault trees and instructions were read aloud.

Results

Focusing subjects' attention on the "other" branch of the unpruned tree had little effect on the unpruned tree group. The mean proportion of starting failures assigned to "other" was actually slightly smaller than the proportion in Experiment 1 (.067 vs. .078), $t(121) = .71$. It did, however, increase by about 50% the proportion of problems that subjects who saw the pruned trees attributed to "other." For pruned tree Group 1, the proportion attributed to "other" increased from .140 to .217, $t(49) = 1.88$; for pruned tree Group 2, from .227 to .346, $t(58) = 1.81$. In both cases, though, the proportion attributed to "other" was still much less than it should have been. These results are shown in Table 2.

Subjects who received the unpruned tree with its full panoply of problems thought that the absolute likelihood of a starting failure was greater than did subjects who saw the pruned trees (Mann-Whitney U test, zs = 2.42 and 1.77, for others' cars and own car, respectively). About 90% of the subjects who saw the unpruned tree thought that starting failure was more likely than a flat tire, compared with about 75% of subjects who saw the pruned tree ($p <$ .10 for difference in proportions, both for others' car and for own car). Whereas the median pruned tree subject thought that a starting failure was 5 times as likely as a flat (for either kind of car), subjects who saw the unpruned tree thought that a starting failure was 20 times more likely for a randomly selected car and 60 times more likely for their own car (Mann-Whitney U test, zs = 2.26 and 4.26, respectively). Seeing the unpruned tree apparently made a starting failure seem more likely.

Discussion

Focusing subjects' attention on what is missing improved their awareness, but only partially. Because it did not lead to an exaggerated estimate of what was missing, such focusing would appear to be a valuable procedure whenever confronted with a fault

Table 2
Effects of Focusing Subjects' Attention on "All Other Problems"

		Proportion of problems attributed to "other"		Frequency of starting failures					
				Other's car			Own car		
				More likely than a flat			More likely than a flat		
Group	*n*	Observed	Should be	% Subjects	*Mdn* response	*Mdn* out of 1,000	% Subjects	*Mdn* response	*Mdn* out of 1,000
Experiment 1									
Unpruned tree	93	.078				75			
Pruned tree 1	29	.140	.468			59			
Pruned tree 2	26	.227	.611			76			
Experiment 2									
Unpruned tree	30	.067		88.9	20	55	90.4	60	47
Pruned tree 1	22	.217	.468	75.0	3	25	83.3	5	35
Pruned tree 2	34	.346	.611	76.5	10	21	72.7	10	20

Note. Subjects in Experiment 2 were instructed to focus on the other category. Data from Experiment 1 are included here for comparison.

tree or a similar representation. It remains to be seen how fault tree designers and evaluators can be brought to a fuller appreciation of how adequate a problem presentation is, particularly when the missing elements are ones of which they are totally or partially unaware. Perhaps a useful rule of thumb would be: The proportion of missing sources of trouble is proportional to the number of things I can think of that are missing multiplied by a measure of my general familiarity with the system (where greater familiarity is assigned a lower score). In other words, if I don't know much but still can detect something missing, then this fault tree representation is quite incomplete.

Experiment 3: Level of Detail

Another discretionary decision faced by someone designing a fault tree for public display is how much detail to present. A designer must consider whether additional detail serves to inform the viewer or leads to confusion, feelings of incompetence, or undue apprehension over the large number of sources of trouble presented. Experiments 1 and 2 showed that omitted branches are essentially out of mind. Experiments 3 and 4 examined the effect on problem evaluation of exposing more or less detail for the displayed branches.

Method

The car-won't-start fault tree was presented to four different groups of subjects with three different levels of detail, Levels 1 and 3 described above (system names only and full detail) and another called *Level 2/plus*. This intermediate level included everything in Level 2 along with full (Level 3) detail for just one branch. Subjects receiving such a fault tree were told that a similarly detailed analysis could be performed for other branches. One group at Level 2/plus received full detail on the fuel system; another received full detail on the battery system.

One hundred ten new subjects were recruited as in Experiments 1 and 2. Questionnaires were distributed in a group setting, and subjects read them at their own pace.

In this and all subsequent experiments, the proportion estimation task of Experiment 1 (subjects gave proportions to all Level 1 categories) and the four final questions from Experiment 2 (number of failures out of 1,000 and comparison with a flat tire for both others' cars and own car) were used.

Results

The upper half of Table 3 presents results from Experiment 3, along with the results for the corresponding group given Level 2 in Experiment 1. The effects of the various manipulations of detail seem quite modest.

One might expect that when just one system receives more detail than the others, as in the two Level 2/plus groups, that system's proportion of problems would be elevated. However, neither battery nor fuel, when given a greater detail, received a proportion significantly ($\alpha = .05$) larger than these systems received in the other groups.

Similarly, one might expect a lower proportion of problems attributed to mischievous acts at Level 3, when it is the only branch without additional detail (see Figure 1), than for the other detail levels. No such effect was found.

The battery and fuel systems, which had more detail at Level 3 than did the other systems, were judged more likely by that group than by other groups (for battery, $t = 2.95$ compared with Level 2 and $t = .86$ compared with Level 1; for fuel, $t = 1.14$ compared with Level 2 and $t = 2.49$ compared with Level 1; $df = 70$ in all tests).

Finally, the proportion of problems attributed to all other problems decreased monotonically as level of detail increased, for linear trend, $F(1, 141) = 7.40$, $p < .01$.

Despite the expectation that the probability of a starting failure would increase as the level of detail increased, none of the four measures of this probability showed this effect. Indeed, the only such measure, shown in Table 3, starting failures out of 1,000 for others' cars, showed a nonsignificant trend in the opposite direction.

Experiment 4

Before concluding that detail has only a modest effect, we explored the possibility that mode of presentation mattered here

Table 3
Effect of Presenting Varying Degrees of Detail

		Mean proportion of starting failures attributed to							*Mdn* no. starting failures in 1,000 (others' cars)
Level of detail	*n*	Battery	Starting	Fuel	Ignition	Engine	Mischief	Other	
		Experiment 3—Self-paced							
System names alone (Level 1)	35	.303	.162	.149	.156	.062	.052	.116	112
Minimal detail (Level 2)[a]	35	.231	.205	.194	.151	.083	.046	.092	100
One branch with full detail									
Battery	19	.284[b]	.184	.182	.180	.066	.039	.065	100
Fuel	19	.224	.256	.174[b]	.140	.067	.051	.088	100
Full Detail (Level 3)	37	.337	.126	.232	.125	.061	.057	.062	70
		Experiment 4—Read aloud							
System names alone (Level 1)	35	.330	.152	.136	.172	.067	.060	.083	27
Minimal Detail (Level 2)[c]	58	.284	.189	.192	.139	.072	.054	.070	28
One branch with full detail									
Battery	22	.347[b]	.185	.175	.124	.052	.051	.066	50
Starting	26	.260	.241[b]	.120	.144	.076	.062	.095	50
Fuel	23	.304	.189	.184[b]	.127	.067	.054	.076	90
Full detail (Level 3)	67	.294	.198	.186	.151	.080	.038	.053	50

[a] Unpruned self-paced group from Experiment 1.
[b] Single branch presented in full detail.
[c] Unpruned read-aloud group from Experiment 1.

(although it had not mattered in Experiment 1). We speculated that subjects presented the rather voluminous fault trees for Levels 2/plus and 3 may have skimmed over them, not attending to the greater detail they offered. Experiment 4 replicated Experiment 3 with fault trees read aloud to 173 new subjects. A third Level 2/plus group was added, giving full detail on the starting system.

Results

The lower portion of Table 3 presents the results from Experiment 4, along with the appropriate Level 2 results from Experiment 1. With Levels 1 and 2, the differences between Experiments 3 and 4 were negligible. The mean absolute differences in proportions for the seven branches were .016 and .018, respectively. With Level 2/plus, there was somewhat more of a difference. As a highlighted branch, the battery system increased from .284 in Experiment 3 to .347 in Experiment 4, $t(39) = 1.33$; the fuel system, however, changed little (.174 vs. .184), $t(40) = .34$.

Unlike Experiment 3, here the battery system received a somewhat (but nor enormously) greater share of attributed problems when it was highlighted (Level 2/plus) than in all other conditions. So did the starting system. However, the fuel system again failed to show this effect.

Unlike Experiment 3, mischievous acts did show a significantly smaller proportion of problems at Level 3, where it is the only branch lacking full detail ($z = 2.95$ compared with Level 1; $z = 2.17$ compared with Level 2).

Again in contrast to Experiment 3, the battery and fuel systems, the ones with the greatest detail, were not more likely in Level 3, where this detail is most apparent, than in other levels.

The proportion of problems attributed to all other problems in Experiment 4 showed the same tendency to decrease as detail increased, as found in Experiment 3, except for one reversal: The weighted mean across the three Level 2/plus groups was .080, rather than a value between .053 and .070. Despite the one reversal, the linear trend in a one-way analysis of variance was significant, $F(1, 227) = 4.44$, $p < .05$.

In Experiment 4, the median starting failure rate for others' cars was similar for Levels 1 and 2 (Mann-Whitney U test, $z = .56$) but somewhat higher for Levels 2/plus and 3 than for Levels 1 and 2 (Mann-Whitney U test, $z = 1.83$). No such differences, even marginally significant, were observed for "one's own car" (not shown), suggesting that subjects' judgments regarding their own cars were less susceptible to this manipulation. Subjects in Experiment 4 also showed some effect due to detail on the likelihood of starting failure compared with that of a flat tire for both their own and others' cars. For Level 1, roughly 75% of subjects thought that starting failure was more likely; for Levels 2, 2/plus, and 3, approximately 90% thought so. The difference in proportions was significant ($z = 2.12$ for others' cars; $z = 2.47$ for own car).

Discussion

Compared with the enormous availability effect found in Experiments 1 and 2, the present results are modest. Only one of several hypothesized effects due to variation in detail was found in both Experiments 3 and 4: The proportion of problems attributed to all other problems decreased as detail increased. Further research with greater power is needed to explore the other effects that were found in one but not in both experiments. Even if real, these possible effects are probably small.

One must conclude that amounts of detail did not produce large changes in people's perceptions. The mere mention of a branch (Level 1) appeared to allow subjects to make a fairly accurate estimate of how troublesome that branch would look when fully detailed. Level 2/plus subjects were somewhat (although not completely) successful in compensating for the detail missing on the more minimally presented branches. The relatively small effect of increased detail on estimated starting failure rate suggests that subjects given little detail realize their own ignorance.

Experiment 5: Splitting and Fusing Branches

A third area in which the designer of a fault tree must make discretionary decisions is in organizing the various sources of trouble into branches. Although functional relationships introduce some constraints, the designer must make decisions like whether "disruptions of wiring" belongs under mischievous acts or "ignition system defective." A common dilemma is broad versus narrow categorization. Starting system defective and ignition system defective could be lumped together. Ignition system defective could be split into ignition system defective (Items 1 and 3 in Figure 1) and "distribution system defective" (Items 2 and 4). Experiment 5 examines the effect on fault tree evaluation of the way in which a fixed amount of information is organized, specifically, whether broader or narrower categorization is used.

Method

The effect of breadth of categorization on fault tree evaluation was studied by creating four new versions of the Level 2 fault tree (Figure 1). Two were created by splitting different existing branches into two branches containing (between them) the same information. In one, ignition system defective was split into ignition system defective with Items 1 and 3 (coil faulty and spark plugs defective) and distribution system defective, with Items 2 and 4 (distributor faulty and defective wiring between components); in the second version, "fuel system defective" was split into fuel system defective (Items 1 and 2) and "carburetion defective" (Items 3 and 4). Thus, these two trees each had eight branches (counting other). Two additional trees were created by fusing two branches of the full tree. In one, starting system defective and ignition system defective were combined into one branch; in the second, fuel system defective and "other engine problems" were combined. Thus these fused versions each had six branches (counting other).

One hundred fifteen subjects were recruited as before. Each subject saw only one tree. The tasks for the subjects were the same as in Experiments 3 and 4. The trees were read aloud.

Results

Table 4 presents the proportion of starting failures attributed to each of the manipulated branches, along with comparable data from Experiment 1. In every case, a set of problems was perceived as more important when it was presented as two branches than when presented as one. The mean increase over the four cases was .066; in general, a set of problems was attributed about a third greater portion of the total number of starting failures when presented as two branches. The number of subjects given eight-branch trees who assigned the two new branches a combined proportion higher than the mean assigned by subjects in Experiment 1 to the comparable single branch was 20 out of 27 (sign test, $p =$ .010) for the fuel-system-split group and 19 out of 26 ($p = .014$) for the ignition-system-split group. The number of subjects who assigned a proportion to the new fused branch lower than the corresponding mean sum in Experiment 1 was 24 out of 33 ($p =$ .007) for the starting-ignition-combined group and 23 out of 29 ($p = .001$) for the fuel-engine-combined group.

There were no differences in either the relative (compared to a flat tire) or absolute likelihood of a starting failure for own or others' cars for subjects seeing the information presented in six- or eight-branch fault trees (not shown in table).

Discussion

Experiment 5 showed that the more pieces into which a system of failure pathways is organized, the more important that system seems. One possible explanation is that people tend to assign some minimum probability to any category with which they are faced. A branch that is split in two receives two portions of this minimum probability, either because the smallest nonzero estimate allowed by the response mode was .01 or because of an assumption that "if they decided to include this branch, it must have some minimal import." An indirect way to evaluate this hypothesis is to estimate what proportion of problems is attributed to a category of "minimal import." For the 93 subjects in Experiment 1 who saw the unpruned Level 2 tree, the mean proportion of

Table 4
Experiment 5: Effects of Splitting and Fusing Branches

	Separate		Together	
Problem	*M* proportion of starting failures	Sum	*M* proportion of starting failures	Difference[a]
	Splitting existing branches			
Fuel	.161	.260	.193	.067
Carburetion	.099			
n = 27				
Ignition	.082	.193	.144	.049
Distribution	.111			
n = 26				
	Fusing existing branches			
Starting	.195	.339	.248	.091
Ignition	.144			
n = 33				
Fuel	.193	.269	.213	.056
Other engine	.076			
n = 29				

Note. In all places in which the number of subjects is not indicated, the results are those from the unpruned tree (Level 2) subjects in Experiment 1 (Table 1), for which *n* = 93.
[a] Separate sum minus together.

problems assigned to the least important branch was .033; the mean lowest proportion excluding zero responses was .040. (Only 11% of subjects ever assigned a zero proportion of problems to any branch.) By this criterion, the increased importance (.066) garnered by splitting a branch was about twice that which might be attributed merely to increasing the number of categories.

Experiment 6: Experts

Our discussion so far has been concerned with lay persons' evaluations of fault trees presented to them by technical experts. One might wonder, though, whether the effects we have found also affect the technical experts themselves. Fischhoff (1977) lists many incidents in which experts designing fault trees for important technological systems were apparently unaware of major omissions; they therefore greatly overestimated the exhaustiveness of their own analyses. Experiment 6 examined whether that which is out of sight is also out of mind for technical experts.

Experiment 6 replicated Experiment 1 using as subjects experienced mechanics in Eugene, Oregon. Thirty copies of Pruned Tree 1 (lacking the starting system, ignition system, and mischief branches) and 30 copies of the unpruned tree (Level 2) were distributed to experienced mechanics at six major garages.

In the accompanying letter, they were told:

> We are asking people like yourself who work with cars to give us their opinion about certain types of engine-starting problems. We hope you will agree to read the questionnaire and give us your opinions.
>
> Actually, we are not just interested in automobile engines. We're concerned with all kinds of complex mechanical systems ranging from automobile engines to nuclear reactors. In particular, we're interested in the ways these systems break down and the judgments about these breakdowns made by people who repair these systems. That's why we're asking for your opinions.
>
> If you are willing to participate in this study, please read the instructions on the next page; then fill out the two short questionnaires and mail them to us in the enclosed addressed and stamped envelope. We'll mail you a check for $3.

> There are no "right" answers to these questions. At least we don't know the answers. That's why we're interested in your opinions. Please work alone and don't refer to any books or discuss the questionnaire with anyone before you have completed it.

Otherwise, the questionnaire was like that in Experiment 1. On the final page, subjects were asked to indicate (a) the rate of starting failures per 1,000 attempts (for drivers in the United States); (b) the number of years they had made all or part of their living working with cars; and (c) how knowledgeable they were about automobile engine problems compared to other mechanics: below average, average, above average, or much above average.

Results

Twenty-nine of the 60 questionnaires were returned, 16 from the pruned tree group and 13 from the unpruned tree group. Two mechanics rated themselves as much above average in knowledge, 18 as above average, 8 as average, and 1 as below average. They had from 2 to 43 years of experience, with the mean for the unpruned tree group being 12.2 years, and for the pruned tree group, 19.8 years.

Table 5 presents the mean proportion of starting failures attributed to each branch by subjects in the present experiment as well as the unpruned tree results from Experiment 1 for the sake of comparison. The experts thought that battery and ignition were more serious problems than did the regular subjects, and that the starting and fuel system problems were less important. They more or less agreed about the completeness of the tree, assigning .060 to "other."

The combined proportion of problems attributed to the branches deleted from the pruned tree (plus "other") was .441. Pruned tree subjects assigned to "other" a mean proportion of only .215 (Mann-Whitney *U* test, $z = 2.87$). The respective median responses were .47 and .16.

Within the pruned tree group, the two subjects who rated themselves as much above average in knowledge assigned a mean to "other" of .075, the nine above-average mechanics assigned .278; and the five average ones, .158. There was a rank order correlation (tau) of .058 between number of years of experience and proportion assigned to "other." Thus, neither self-rated degree of knowledge nor actual experience had any systematic relation to ability to detect what was missing from the tree.

General Discussion

The most dramatic result of these studies was subjects' inability to appreciate how much had been omitted from the pruned fault trees. Exaggeration of fault tree completeness was found with both "regular" subjects and experienced mechanics. The

Table 5
Experiment 6: Judgment of Experts

		Mean proportion of starting failures attributed to							
Group	*n*	Battery	Starting system	Fuel system	Ignition system	Engine	Mischief	Other	*Mdn* no. starting failures in 1,000
Unpruned tree ordinary subjects	93	.264	.195	.193	.144	.076	.051	.078	75
Unpruned tree experts	13	.410	.108	.096	.248	.051	.025	.060	20
Pruned tree 1 experts	16	.483	—	.229	—	.073	—	.215[a]	100

[a] Should be .441.

fact that omission of major branches triggered only minimal awareness of the inadequacies of the pruned tree lent strong support to the availability hypothesis over the credibility hypothesis, as did the modest improvement when subjects' attention was focused on completeness.

How might things that are out of sight also be out of mind? One obvious reason is ignorance. There is no way to consider something that one has never heard of and that is not mentioned. In a discussion of the omissions that seem to plague technical experts performing formal risk assessments, Fischhoff (1977) suggested several other reasons: (a) failure to consider the imaginative ways in which human error can mess up a system (e.g., the Brown's Ferry fire in which the world's largest nuclear power plant almost melted down due to a technician checking for an air leak with a candle in direct violation of standard operating procedure); (b) insensitivity to the assumptions an analysis makes about constancies in the world in which the system is embedded (e.g., no major changes in government regulatory policy); (c) overconfidence in current scientific and technological knowledge (i.e., assuming that there are no new chemical, physical, biological, or psychological effects to be discovered); and (d) failure to see how the system functions as a whole (e.g., a system may fail because a backup component has been removed for routine maintenance).

Although similar problems seem likely to afflict the designers and viewers of fault trees, this list is probably incomplete. Expanding and validating such a list is important not only for our understanding of how people conceptualize complex fallible systems but also for helping them better describe and comprehend such systems. Focusing on "other" was somewhat, but only partially, successful in helping people appraise the completeness of a fault tree. Improved understanding of the reasons for omissions will help in designing better "debiasing" procedures and ways to approach an evaluation task that provide a more veridical perspective. When such "cognitive engineering" seems ineffective, people may need rules of thumb like the suggested rule for estimating how many sources of trouble one cannot think of on the basis of how many one can think of and one's familiarity with the problem area.

Because of the importance of intelligent public participation in debates about technology (Casper, 1976; Slovic, Fischhoff, & Lichtenstein, 1976), these are critical issues to which psychologists might address themselves. Such research does, however, raise serious ethical questions because of the possibility that the results will disclose ways in which public opinion can be manipulated. For example, Experiments 1 and 6 suggest that one can get the public to focus on those issues one thinks are important by never mentioning other issues. Even our "debiasing procedure" (focus on "other") in Experiment 2 was a form of manipulation, changing people's perceptions from what they would otherwise be. (Although we believe that this change is for the better, the point is moot as long as we do not know the proportion of problems in fact due to each branch.) Perhaps the prime responsibility of the discoverer of such effects is to ensure that they receive the broadest possible dissemination, so that both potential manipulators and the potentially manipulated are aware of them. It may turn out that we psychologists are merely discovering "tricks" that manipulators have known about all along. Such research can also suggest aspects of presentation that may have relatively little effect on people's judgments (varying level of detail), attempted manipulations that can be readily overcome (providing more detail for one branch), and manipulations whose effects, although consistent, may be too small to have applied implications.

Aside from calling for obviously needed constructive replications (different trees, different subjects, different experts, etc.), these results suggest a variety of questions for future study: Would credibility be a stronger effect than availability if pruned fault trees were presented by distrusted technocrats rather than moderately trusted experimental psychologists? What happens when people construct their own fault trees? What happens when the detailed items from a fault tree are listed but not organized into categories? What happens with trees dealing with more technical areas or with nontechnical areas for which the assessors' feelings of competence vary? What happens when minor rather than major items are deleted? What if an omitted item is somehow brought to the subjects' attention (as often happens in public debates)? What individual differences, if any, interact with these aspects of presentation in affecting evaluations? (As mentioned, the obvious covariate of years of experience was uncorrelated with the mechanics' awareness of how much had been deleted from the pruned tree.) What if detail is manipulated not by adding facts but by adding flourishes, by "fleshing out" possible sources of trouble with vividly detailed scenarios describing just how they might happen? Why were subjects' perceptions of the overall starting failure rate relatively impervious to our manipulations? Was it merely a problem of measurement?

These studies have supported the notion that availability affects frequency estimates and have helped clarify how availability mechanisms work. Failure to mention a possibility had a major effect, increasing the detail of what was mentioned had a small effect, and the packaging of what is mentioned also made a difference. Experiment 2 showed, as did Beyth-Marom and Fischhoff (1977), that forcing people to work their memories (or imaginations) harder can improve their likelihood judgments somewhat. Clearly, much more work is needed on the retrieval, perception, and representation of frequency information.

Reference Note

1. Bryan, W. B. *Testimony before the Subcommittee on State Energy Policy*. Committee on Planning, Land Use, and Energy, California State Assembly, February 1974.

References

Atomic Energy Commission. Reactor safety study: *An assessment of accident risks in U.S. commercial power plants (WASH-1400)*. Washington, D.C.: Author, 1975.

Barrager, S. M., Judd, B. R., & North, D. W. *The economic and social costs of coal and nuclear electric generation.* Washington, D.C.: National Science Foundation, 1976.

Beyth-Marom, R., & Fischhoff, B. Direct measures of availability and judgments of category frequency. *Bulletin of the Psychonomic Society,* 1977, *9,* 236–238.

Birnbaum, M. H., Wong, R., & Wong, L. K. Combining information from sources that vary in credibility. *Memory & Cognition,* 1976, *4,* 330–336.

Casper, B. M. Technology policy and democracy. *Science,* 1976, *194,* 29–35.

Fischhoff, B. Cost benefit analysis and the art of motorcycle maintenance. *Policy Sciences,* 1977, *8,* 177–202.

Green, A. E., & Bourne, A. J. *Reliability technology.* New York: Wiley-Interscience, 1972.

Howell, W. C. Representation of frequency in memory. *Psychological Bulletin,* 1973, *80,* 44–53.

Kendall, H. W. *Nuclear power risks: A review of report of APS society's study group on light water reactor safety.* Cambridge, Mass.: Union of Concerned Scientists, 1975.

Newell, A., & Simon, H. A. *Human problem solving.* Englewood Cliffs, N.J.: Prentice-Hall, 1972.

Settle, R. B., & Golden, L. L. Attribution theory and advertiser credibility. *Journal of Marketing Research,* 1974, *11,* 181–185.

Slovic, P., Fischhoff, B., & Lichtenstein, S. Cognitive processes and societal risk taking. In J. S. Carroll & J. W. Payne (Eds.), *Cognition and social behavior.* Potomac, Md.: Erlbaum, 1976.

Tversky, A., & Kahneman, D. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology,* 1973, *4,* 207–232.

Received June 2, 1977 ■